

- 89 -

- (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AATGCGGCCA CTCCAAGGAG GCCGGGAGGA TTGTGGGAGG CCAAGACACC CAGGAAGGAC	60
GCTGGCCGTG GCAGGTGGC CTGTGGTTGA CCTCAGTGGG GCATGTATGT GGGGGCTCCC	120
TCATCCACCC ACGCTGGGTG CTCACAGCCG CCCACTGCTT CCTGAGGTCT GAGGATCCCG	180
GGCTCTACCA TGTTAAAGTC GGAGGGCTGA CACCCTCACT TTCAGAGCCC CACTCGGCCT	240
TGGTGGCTGT GAGGAGGCTC CTGGTCCACT CCTCATACCA TGGGACCACC ACCAGCGGGG	300
ACATTGCCCT GATGGAGCTG GACTCCCCCT TGCAGGCCTC CCAGTTCAGC CCCATCTGCC	360
TCCCAGGACC CCAGACCCCC CTCGCCATTG GGACCGTGTG CTGGGTAAAC GGGCTGGGGG	420
TCCACTCAGG AGAGGCCCTG GCGAGTGTC TTCAGGAGGT GGCTGTGCCC CTCCTGGACT	480
CGAACATGTG TGAGCTGATG TACCACCTAG GAGAGCCCAG CCTGGCTGGC CAGCGCCTCA	540
TCCAGGACGA CATGCTCTGT GCTGGCTCTG TCAGGGCAA GAAAGACTCC TGCCAGGGTG	600
ACTCCGGGGG GCCGCTGGTC TGCCCCATCA ATGATACGTG GATCCAGGCC GGCATTGTGA	660
GCTGGGGATT CCGCTGTGCC CGGCCTTTCC GGCCTGGTGT CTACACCCAG GTGCTAAGCT	720
ACACAGACTG GATTCAGAGA ACCCTGGCTG AATCTCACTC AGGCATGTCT GGGGCCCCGCC	780
CAGGTGCCCC AGGATCCCAC TCAGGCACCT CCAGATCCCA CCCAGTGCTG CTGCTTGAGC	840
TGTTGACCGT ATGCTTGCTT GGGTCCCTGT GAACCATGAG CCATGGAGTC CGGGATCCCC	900
TTTCTGGTAG GATTGATGGA ATCTAATAAT AAA	933

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- 90 -

- (A) LENGTH: 980 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCTGTGGTCG CCCAGGATG CTGAACCGAA TGGTGGGCGG GCAGGACACG CAGGAGGGCG	60
AGTGGCCCTG GCAAGTCAGC ATCCAGCGCA ACGGAAGCCA CTTCTGCGGG GGCAGCCTCA	120
TCGCGGAGCA GTGGGTCTG ACGGCTGCGC ACTGCTTCCG CAACACCTCT GAGACGTCCC	180
TGTACCAGGT CCTGCTGGGG GCAAGGCAGC TAGTGCAGCC GGGACCACAC GCTATGTATG	240
CCCGGGTGAG GCAGGTGGAG AGCAACCCCC TGTACCAGGG CACGGCCTCC AGCGTGACG	300
TGGCCCTGGT GGAGCTGGAG GCACCAGTGC CCTTCACCAA TTACATCCTC CCCGTGTGCC	360
TGCCTGACCC CTCGGTGATC TTTGAGACGG GCATGAACTG CTGGGTCACT GGCTGGGGCA	420
GCCCCAGTGA GGAAGACCTC CTGCCCCAAC CGCGGATCCT GCAGAACTC GCTGTGCCCA	480
TCATCGACAC ACCCAAGTGC AACCTGCTCT ACAGCAAAGA CACCGAGTTT GGC'TACCAAC	540
CCAAAACCAT CAAGAATGAC ATGCTGTGCG CCGGCTTCGA GGAGGGCAAG AAGGATGCCT	600
GCAAGGGCGA CTCGGGCGGC CCCCTGGTGT GCCTCGTGGG TCAGTCGTGG CTGCAGGCGG	660
GGGTGATCAG CTGGGGTGAG GGCTGTGCCC GCCAGAACCG CCCAGGTGTC TACATCCGTG	720
TCACCGCCCA CCACAACTGG ATCCATCGGA TCATCCCCAA ACTGCAGTTC CAGCCAGCGA	780
GGTTGGGCGG CCAGAACTGA GACCCCCGGG GCCAGGAGCC CCTTGAGCAG AGCTCTGCAC	840
CCAGCCTGCC CGCCACACC ATCCTGCTGG TCCTCCAGC GCTGCTGTTG CACCTGTGAG	900
CCCCACCAGA CTCATTTGTA AATAGCGCTC CTTCTCCCC TCTCAAATAC CCTTATTTTA	960
TTTATGTTTC TCCCAATAAA	980

## CLAIMS:

1. An isolated proteinaceous molecule involved in or associated with regulation of cell activity and/or viability comprising a sequence of amino acids encoded by a nucleotide sequence, at least a portion of which, is capable of being amplified by polymerase chain reaction (PCR) using the following primers:

5' ACAGAATTCTGGGTIGTTIACIGCIGCICAYTG3' [SEQ ID NO:1]; and

5' ACAGAATTCA XIGGICCCIC/GT/AXTCICCC3' [SEQ ID NO:2];

or a complementary form of said primers.

2. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.

3. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.

4. An isolated proteinase molecule according to claim 1 wherein said molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

5. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:3 under low stringency conditions at 42°C.

6. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions at 42°C.
7. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:7 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:7 under low stringency conditions at 42°C.
8. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a kinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:10 or having 50% amino acid similarity thereto.
9. An isolated proteinaceous molecule according to claim 1 wherein said molecule is a kinase comprising an amino acid sequence encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the nucleotide sequence set forth in SEQ ID NO:9 under low stringency conditions at 42°C.
10. An isolated nucleic acid molecule encoding a polypeptide wherein at least a portion of said nucleic acid molecule is capable of being amplified by polymerase chain reaction (PCR) using the following primers:

5' ACAGAATTCTGGGTIGTACIGCIGCICAYTG3' [SEQ ID NO:1]; and

5'ACAGAATTCA XIGGICCCIC/C/GT/AXTCICCC3' [SEQ ID NO:2];

or a complementary form of said primers.

11. An isolated nucleic acid molecule according to claim 10 wherein said polypeptide is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.
12. An isolated nucleic acid molecule according to claim 10 wherein said polypeptide is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.
13. An isolated nucleic acid molecule according to claim 10 wherein said polypeptide is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.
14. An isolated nucleic acid molecule according to claim 10 comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:3 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:3 under low stringency conditions at 42°C.
15. An isolated nucleic acid molecule according to claim 10 comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions at 42°C.
16. An isolated nucleic acid molecule according to claim 10 comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:7 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:7 under low stringency conditions at 42°C.
17. An isolated nucleic acid molecule according to claim 10 wherein said polypeptide is a kinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:10 or having 50% amino acid similarity thereto.

- 94 -

18. An isolated nucleic acid molecule according to claim 17 comprising a sequence of nucleotides encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the nucleotide sequence set forth in SEQ ID NO:9 under low stringency conditions at 42°C.
19. An isolated serine proteinase encoded by a gene proximal to a cluster of genes of a mammalian chromosome.
20. An isolated serine proteinase according to claim 19 wherein the mammalian chromosome is human chromosome 16p13.3 or its equivalent in a non-human species.
21. An isolated serine proteinase according to claim 20 wherein the gene cluster includes at least two genes having the nucleotide sequence as set forth in SEQ ID NO:3 or 5 or 28 or 29 or 30 or a nucleotide sequence having at least 50% similarity to any one of SEQ ID NO:3 or 5 or 28 or 29 or 30 or a nucleotide sequence capable of hybridizing to any one of the sequences under low stringency conditions at 42°C.
22. An isolated serine proteinase according to claim 20 wherein said serine proteinase is a short form of HELA2 having an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.
23. An isolated serine proteinase according to claim 20 wherein said serine proteinase is a long form of HELA2 having an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.
24. An isolated serine proteinase according to claim 22 encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a nucleotide sequence having at least 50% similarity thereto or a sequence capable of hybridizing to SEQ ID NO:3 under low stringency conditions at 42°C.

25. An isolated serine proteinase according to claim 23 encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a sequence capable of hybridizing to SEQ ID NO:5 under low stringency conditions at 42°C.
26. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a serine proteinase and corresponding to a gene proximal to a cluster of genes encoding serine proteinases.
27. An isolated nucleic acid molecule according to claim 26 wherein the gene cluster includes at least two genes having the nucleotide sequence as set forth in SEQ ID NO:3 or 5 or 28 or 29 or 30 or a nucleotide sequence having at least 50% similarity to any one of SEQ ID NO:3 or 5 or 28 or 29 or 30 or a nucleotide sequence capable of hybridizing to any one of the sequences under low stringency conditions at 42°C.
28. An isolated nucleic acid molecule according to claim 25 comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3 or SEQ ID NO:5 or a nucleotide sequence having at least about 50% similarity to either of SEQ ID NO:3 or SEQ ID NO:5 or a nucleotide sequence capable of hybridizing to SEQ ID NO:3 or SEQ ID NO:5 under low stringency conditions at 42°C.
29. An isolated kinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:10 or an amino acid sequence having at least about 50% similarity thereto.
30. An isolated kinase according to claim 29 encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or capable of hybridizing to SEQ ID NO:9 under low stringency conditions at 42°C.
31. A method of regulating cell activity and/or viability said method comprising contacting said cell with an activity and/or viability effective amount of a serine proteinase and/or kinase.

- 96 -

32. A method according to claim 31 wherein the serine proteinase comprises a sequence of amino acids encoded by a nucleotide sequence, at least a portion of which, is capable of being amplified by polymerase chain reaction (PCR) using the following primers:

5' ACAGAATTCTGGGTIGTACIGCIGCICAYTG3' [SEQ ID NO:1]; and

5'ACAGAATTCA XIGGICCCIC/GT/AXTCICCC3' [SEQ ID NO:2];

or a complementary form of said primers.

33. A method according to claim 31 wherein the serine proteinase comprises an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.

34. A method according to claim 31 wherein the serine proteinase comprises an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.

35. A method according to claim 31 wherein the serine proteinase comprises an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

36. A method according to claim 31 wherein the serine proteinase comprises a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:3 under low stringency conditions at 42°C.

37. A method according to claim 31 wherein the serine proteinase comprises a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions at 42°C.



38. A method according to claim 31 wherein the serine proteinase comprises a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:7 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:7 under low stringency conditions at 42°C.
39. A method according to claim 31 wherein the kinase comprises an amino acid sequence substantially as set forth in SEQ ID NO:10 or having 50% amino acid similarity thereto.
40. A method according to claim 31 wherein the kinase comprises an amino acid sequence encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the nucleotide sequence set forth in SEQ ID NO:9 under low stringency conditions at 42°C.
41. A method of modulating fertility in a mammal said method comprising modulating levels of HELA2 wherein increasing levels of HELA2 facilitates sperm maturation and development.
42. A method according to claim 41 wherein fertility is enhanced by introducing recombinant HELA2.
43. A method according to claim 41 wherein fertility is reduced by down regulating expression of the HELA2 gene.
44. A composition comprising a serine proteinase and/or kinase capable of regulating cell activity and/or viability and one or more pharmaceutically acceptable carriers and/or diluents.
45. A composition according to claim 44 wherein the serine proteinase is HELA2 or a functional derivative thereof.
46. An isolated antibody capable of interacting with a proteinaceous molecule involved in or associated with regulation of cell activity and/or viability comprising a sequence of amino acids encoded by a nucleotide sequence, at least a portion of which, is capable of being

- 98 -

amplified by polymerase chain reaction (PCR) using the following primers:

5' ACAGAATTCTGGGTIGTGTACIGCIGCICAYTG3' [SEQ ID NO:1]; and

5' ACAGAATTTCAXIGGICCCIC/GT/AXTCICC3' [SEQ ID NO:2];

or a complementary form of said primers.

47. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.

48. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.

49. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

50. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:3 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:3 under low stringency conditions at 42°C.

51. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions at 42°C.

- 99 -

52. An isolated antibody according to claim 46 wherein said proteinaceous said molecule is a serine proteinase comprising a sequence of amino acids encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:7 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:7 under low stringency conditions at 42°C.

53. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a kinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:10 or having 50% amino acid similarity thereto.

54. An isolated antibody according to claim 46 wherein said proteinaceous molecule is a kinase comprising an amino acid sequence encoded by a nucleotide sequence substantially as set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the nucleotide sequence set forth in SEQ ID NO:9 under low stringency conditions at 42°C.

55. An antagonist or agonist to the isolated proteinaceous molecule according to any one of claims 1 to 9.

56. A method of determining a predisposition for or the presence of a cancer, said method comprising determining the presence of a nucleotide sequence encoding a proteinaceous molecule according to any one of claims 1 to 9.

57. A method according to claim 56 wherein the nucleotide sequence encodes a polypeptide wherein at least a portion of said nucleotide sequence is capable of being amplified by polymerase chain reaction (PCR) using the following primers:

5' ACAGAATTCTGGGTIGTIACIGCIGCICAYTG3' [SEQ ID NO:1]; and

5' ACAGAATTCA XIGGICCCIC/ GT/ AXCICCC3' [SEQ ID NO:2];

or a complementary form of said primers.

58. A method according to claim 57 wherein said nucleotide sequence encodes a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:4 or an amino acid sequence having at least 50% similarity thereto.

59. A method according to claim 57 wherein said nucleotide sequence encodes a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:6 or an amino acid sequence having at least 50% similarity thereto.

60. A method according to claim 57 wherein said nucleotide sequence encodes a serine proteinase comprising an amino acid sequence substantially as set forth in SEQ ID NO:8 or an amino acid sequence having at least about 50% similarity thereto.

61. A method according to claim 57 wherein said nucleotide sequence is as substantially set forth in SEQ ID NO:3 or is a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:3 under low stringency conditions at 42°C.

62. A method according to claim 57 wherein said nucleotide sequence is as substantially set forth in SEQ ID NO:5 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:5 under low stringency conditions at 42°C.

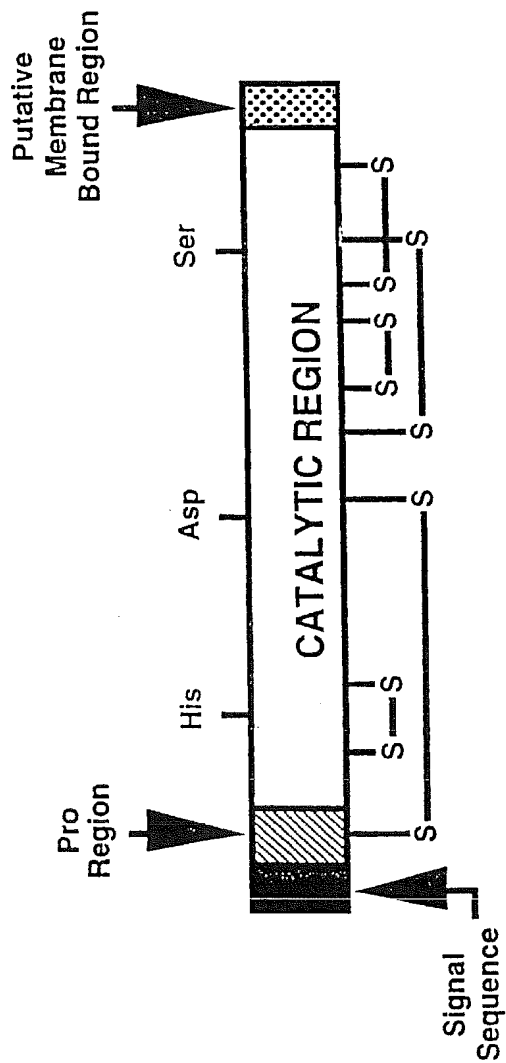
63. A method according to claim 57 wherein said nucleotide sequence is as substantially set forth in SEQ ID NO:7 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the sequence set forth in SEQ ID NO:7 under low stringency conditions at 42°C.

64. A method according to claim 57 wherein said nucleotide sequence is as substantially set forth in SEQ ID NO:10 or having 50% amino acid similarity thereto.

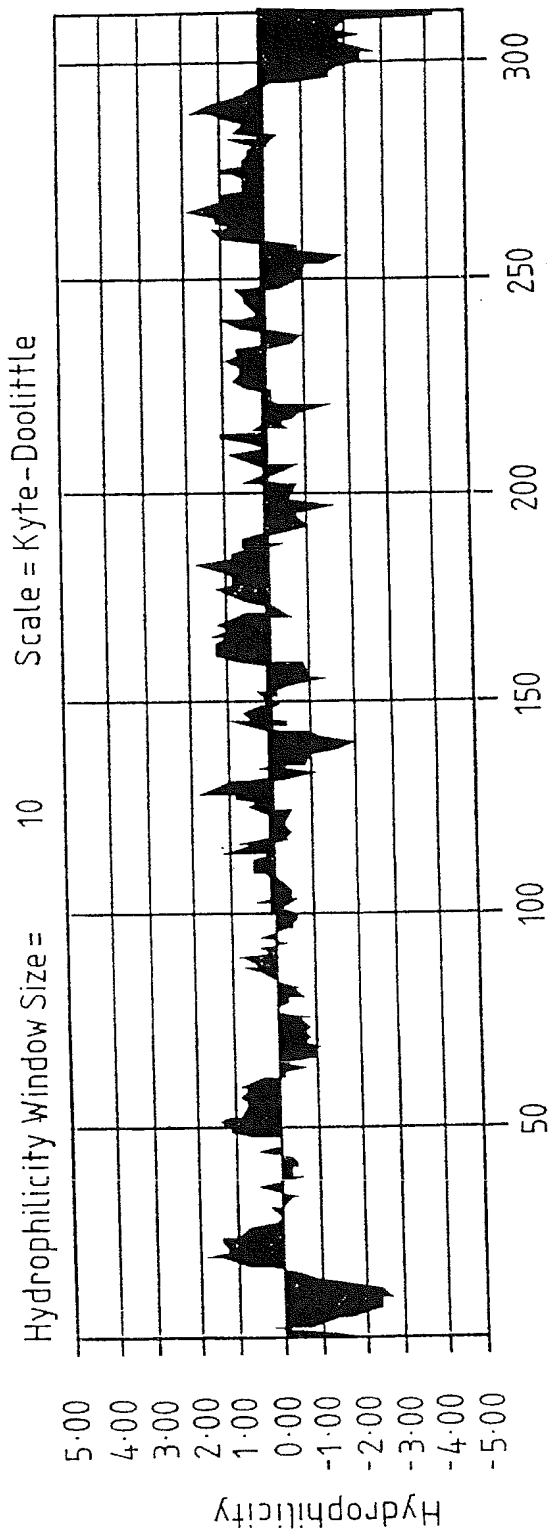
- 101 -

65. A method according to claim 57 wherein said nucleotide sequence is as substantially set forth in SEQ ID NO:9 or a nucleotide sequence having at least 50% similarity thereto or a nucleotide sequence capable of hybridising to the nucleotide sequence set forth in SEQ ID NO:9 under low stringency conditions at 42°C.

1/62

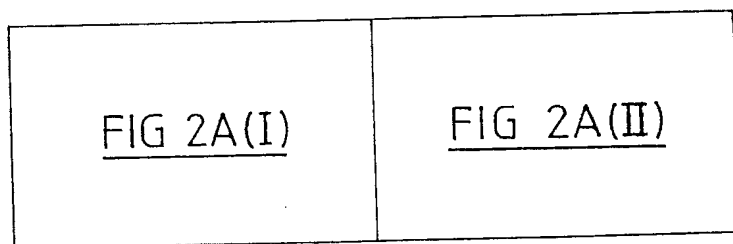


**FIG 1**



2/62

FIG 2A



SUBSTITUTE SHEET (RULE 26)

3/62

## Sequence comparison of HELA2(Testisin) and prostatin

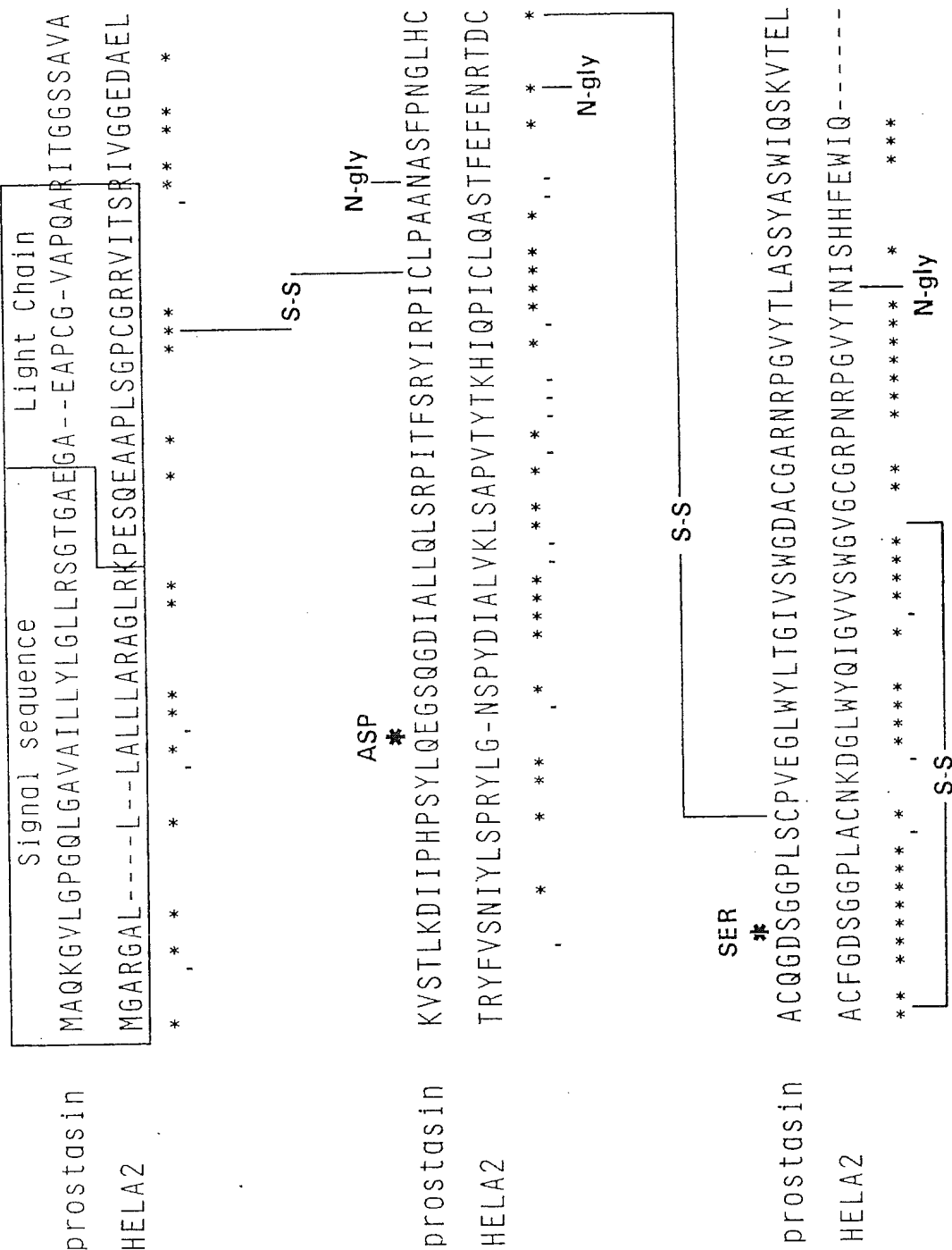


FIG 2A(I)



4/62

Sequence comparison of HELA2(Testisin) and prostatic

prostatic      GQWPWQVSITYEGVHVCGGSLVSEQWVLSAAHCFPSEHK-EAYEVKLG----AHQLDSYSEDA  
 HELA2          GRWPWQGSRLWLWDHSHVCGVSLLSHRWALTAAHCFETDLSDPGWMVQFGQLTMPFWSLQAYY  
 \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*  
 \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*

prostatic      TVTGWHVAPSVSLTPKPLQQLVPLISRETCNCLYNIDAKPEEPHFVQEDMVCAGYVEGGKD  
 HELA2          WVTGWGIKEALPSPHTLQEVQVAIINNSMCNHLFLKYSFRKD--IFG-DMVCAGNAQQGGKD  
 \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*

prostatic      QPRVVPQTQESQPDNSLCGSHLAFSSAPAQGLLR  
 HELA2          --KLMAQSGMSQPD-----PS-----W  
 \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*

Putative Transmembrane Domain  
 PILFLPLGLALGLLSPWL  
 PLLFFPLLWALPLLGPV  
 \* \* \* \* \*      \* \* \* \* \*      \* \* \* \* \*

FIG 2A(II)

5/62

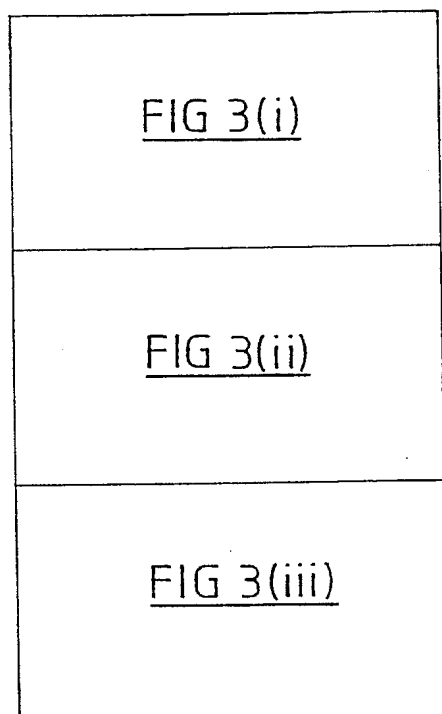
FIG 2B

In vitro transcription /  
translation of HELA2 (Testisin)

SUBSTITUTE SHEET (RULE 26)

6/62

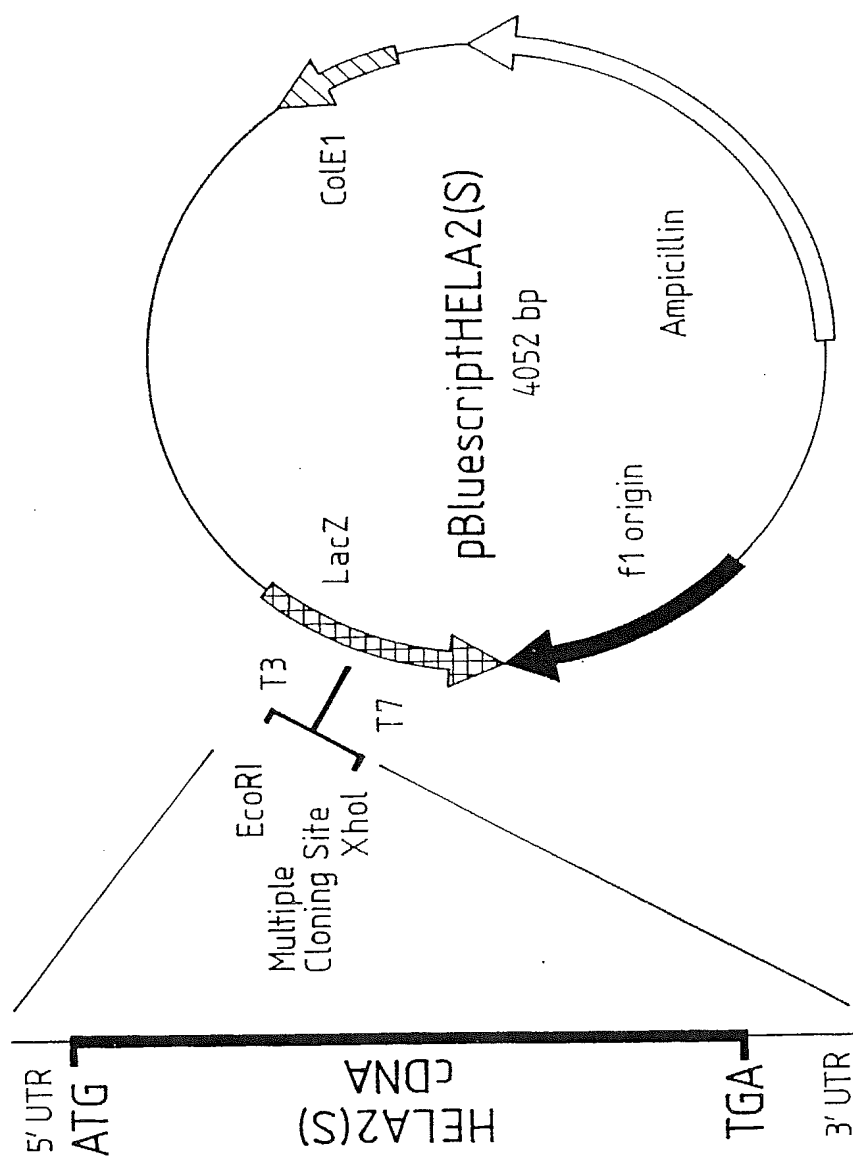
FIG 3



SUBSTITUTE SHEET (RULE 26)

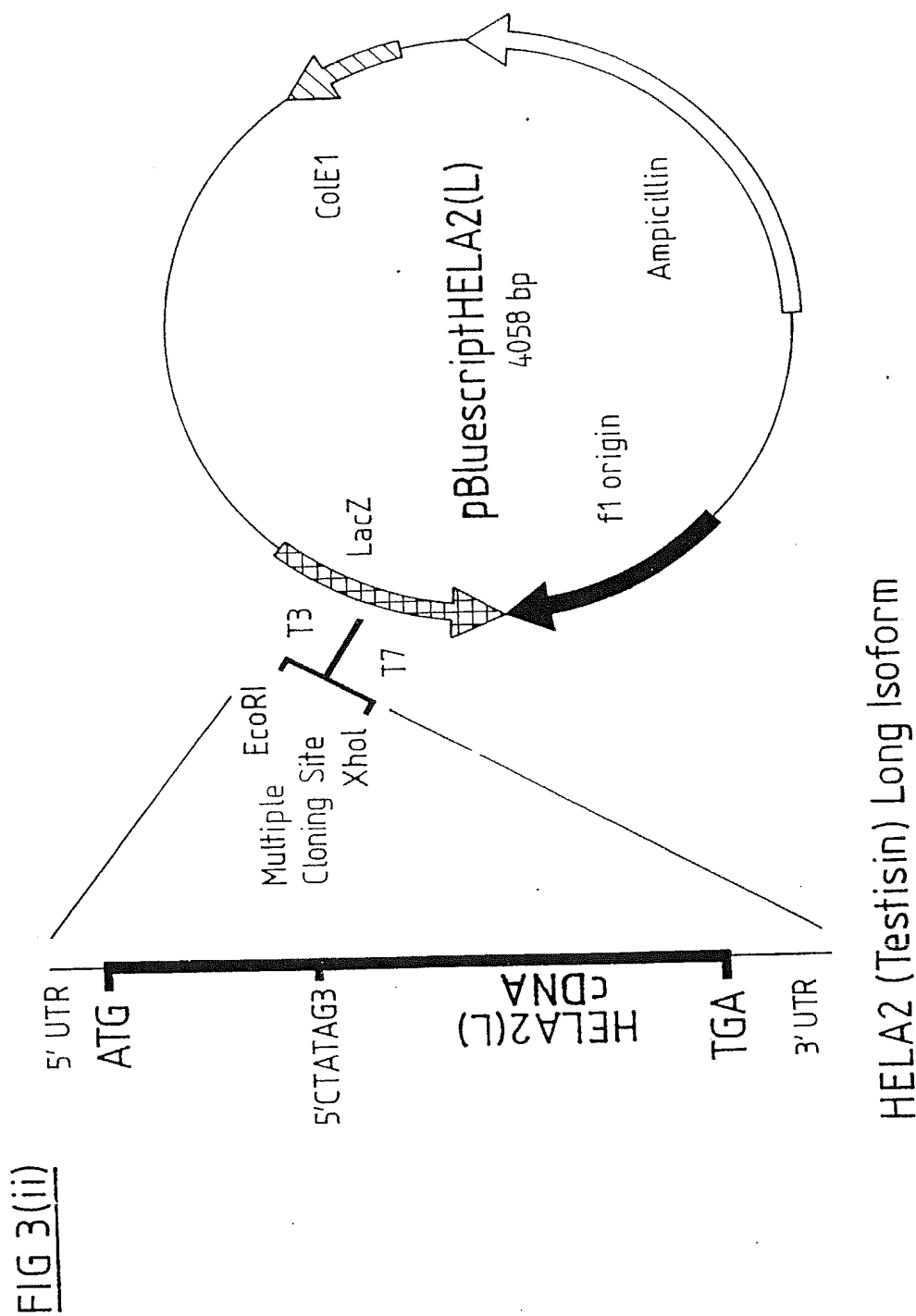
7/62

FIG 3(i)



HELA2 (Testisin) Short Isoform

8/62



SUBSTITUTE SHEET (RULE 26)

9/62

# HELA2 (Testisin) Restriction Enzyme Map

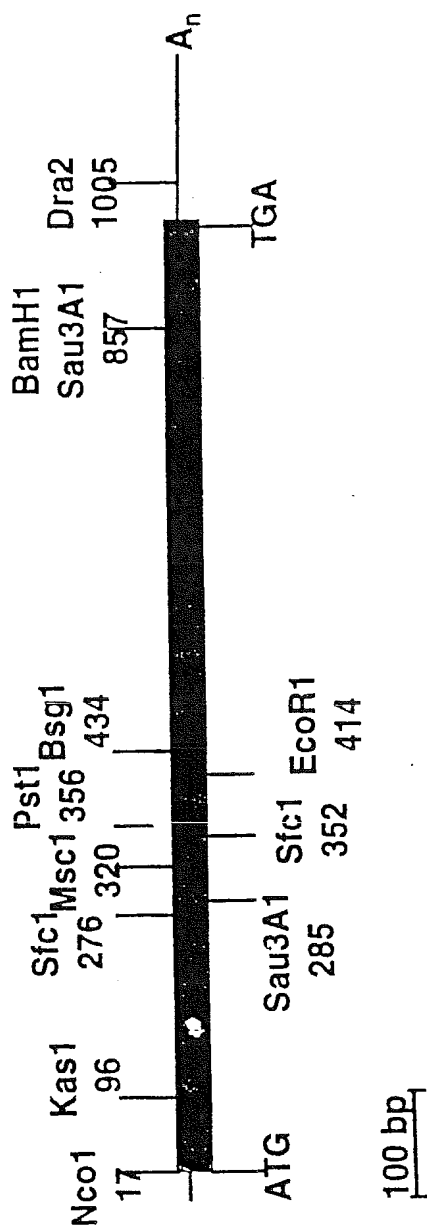
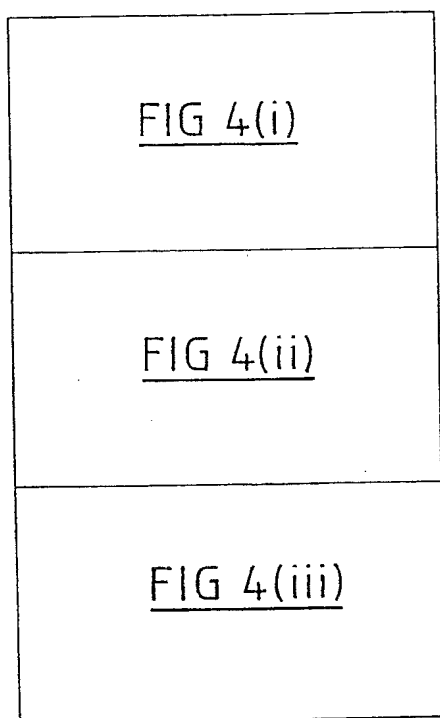


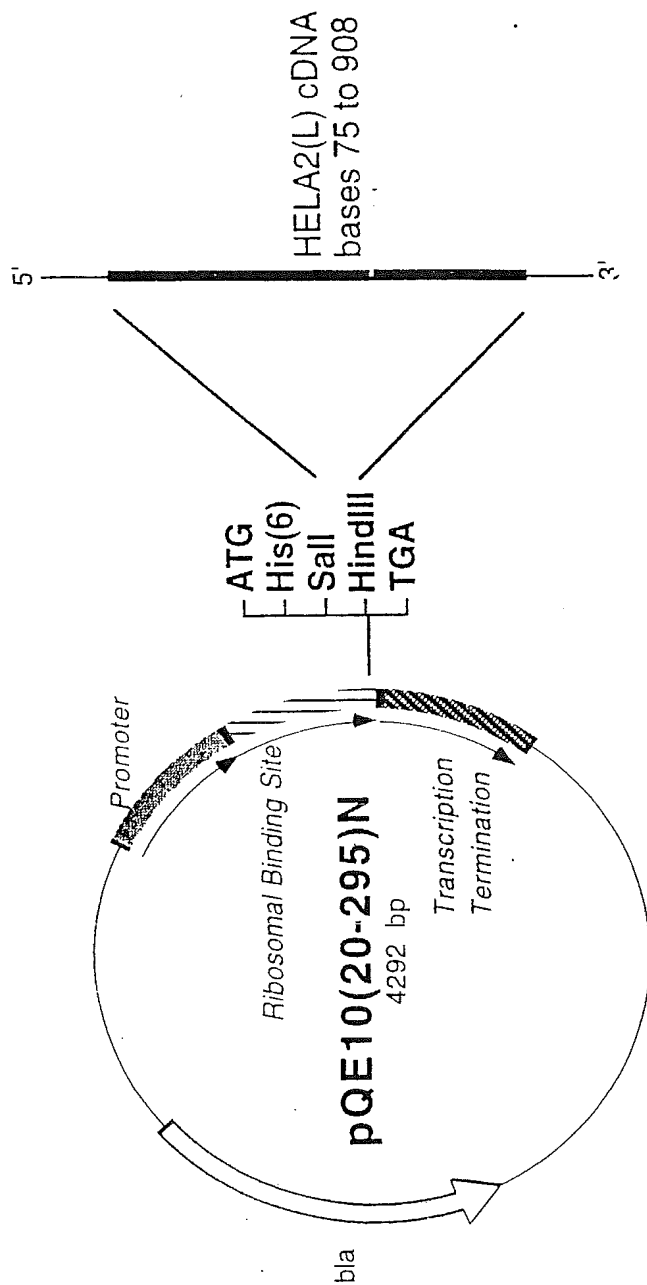
FIG 3(iii)

10/62

FIG 4



11/62



**FIG 4(i)**



12/62

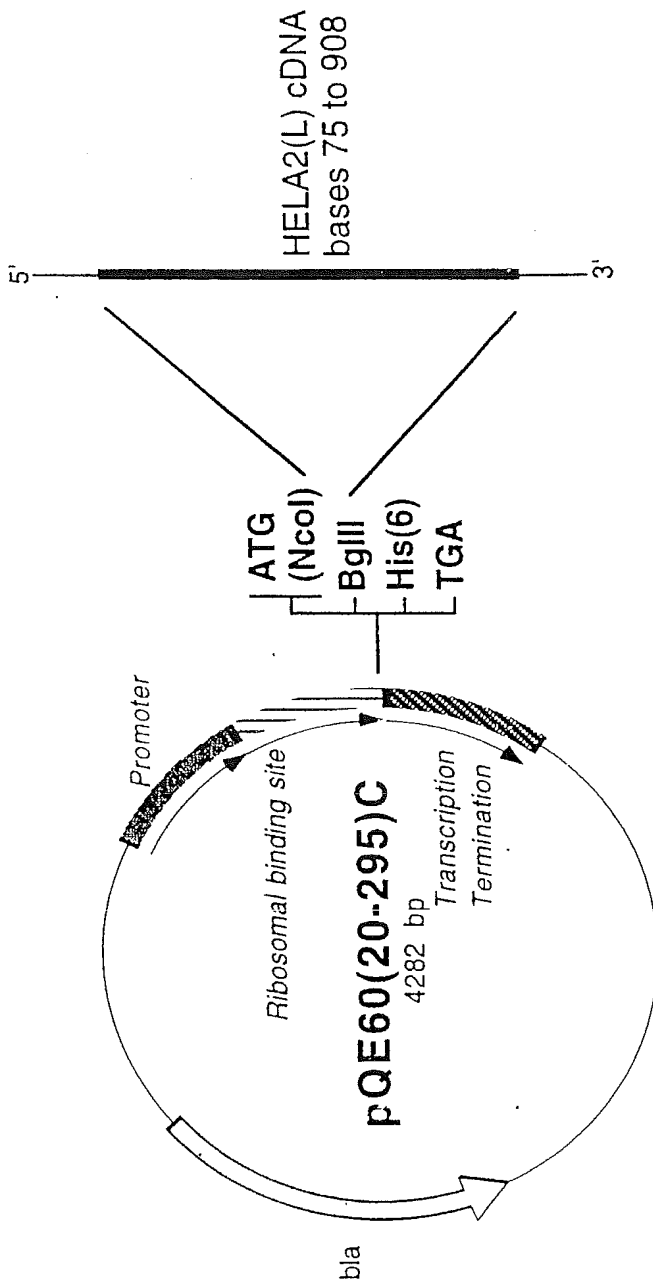


FIG 4(ii)

13/62

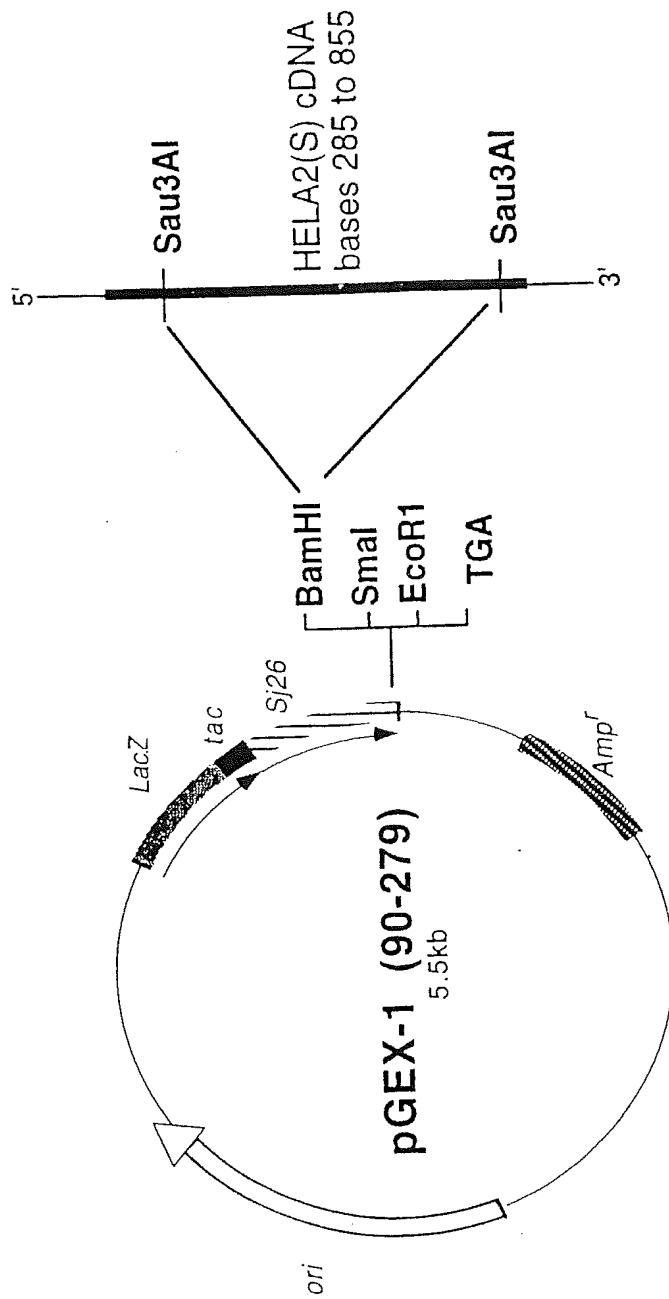
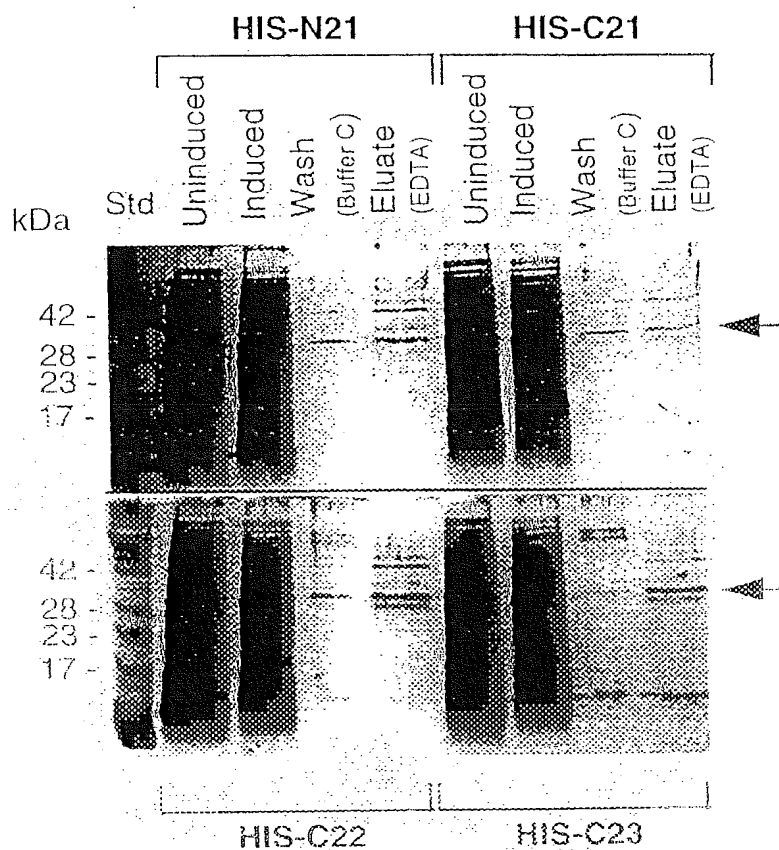
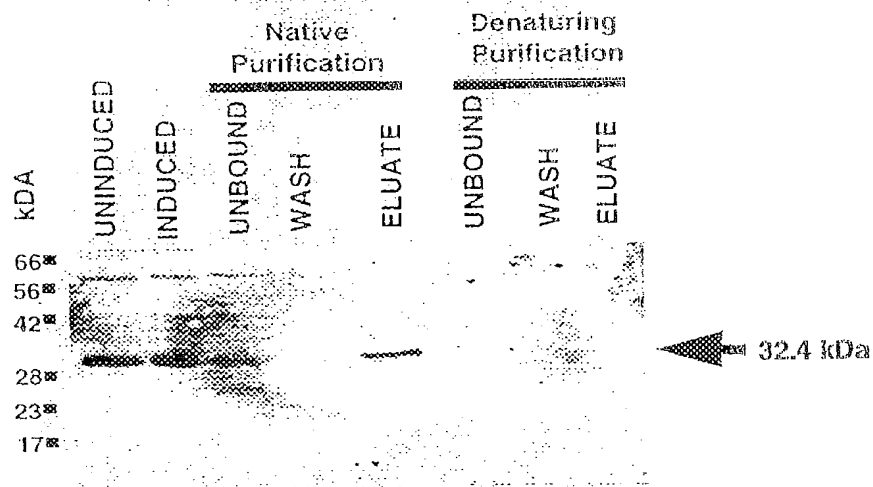


FIG 4(iii)

**FIG 5**

14/62

**A. Expression of recombinant Testisin in *E. coli*.****B. Western blot of recombinant Testisin**

SUBSTITUTE SHEET (RULE 26)

15/62

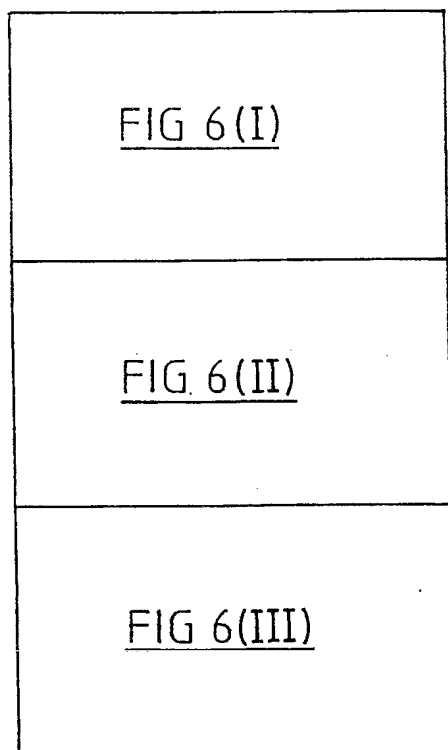


FIG 6

SUBSTITUTE SHEET (RULE 26)

FIGURE 6 (I)

1  
19 ATGGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTCGGGCTGGACTCAGGAAG 20  
M G A R G A L L L A L L A R A G L R K  
79 CCGAGTCGAGAGCGGGCGCGGTTATCAGGACCATGCGGCCGACGGGTATCAGCTCG 40  
P E S Q E A A P L S G P C G R R V I T S  
139 CGCATCGTGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCCGTGGCAGGGGAGCCTGCGC 60  
R I V G G E D A E L G R W P W Q G S L R  
199 CTGTGGATTCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGGCACTCAGCGCG 80  
L W D S H V C G G V S L L S H R W A L T A  
259 GCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGCTCCAGTTT 100  
A H C F E T Y S D L S D P S G W M V Q F  
319 GGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCCTACTACACCCGTTACTTC 120  
G Q L T S M P S F W S L Q A Y Y T R Y F  
379 GTATCGAATATCTATGAGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCCCTTG 140  
V S N I Y L S P R Y L G N S P Y D I A L

16/62

## FIGURE 6 (II)

439 GTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCC  
V K L S A P V T Y T K H I Q P I C L Q A 160

499 TCCACATTTGAGTTGAGAACCGGACAGACTGCTGGTGACTGGCTGGGGTACATCAAA  
S T F E F E N R T D C W V T G W G Y I K 180

559 GAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTCAGGTCGCCATCATAAAC  
E D E A L P S P H T L Q E V Q V A I I N 200

619 AACTCTATGTGCAACCACCTCTTCCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGAC  
N S M C N H L F L K Y S F R K D I F G D 220

679 ATGGTTTGTCTGGCAATGCCCAAGCGGGAAGGATGCCCTTCCGTGACTCAGGTGGA  
M V C A G N A Q G G K D A C F G D S G G 240

739 CCCTTGGCCTGTAAACAAGAAATGGACTGTGTATCAGATTGGAGTCGTGAGCTGGGGAGTG  
P L A C N K N G L W Y Q I G V V S W G V 260

799 GGCTGTGGTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACCTTGAGTGG  
G C G R P N R P G V Y T N I S H H F E W 280

18/62

**FIGURE 6 (III)**

```
859 ATCCAGAAAGCTGATGGCCCAAGAGTGGCATGTCCAGCCAGACCCCTCCTGGCCGCTACTC
    I Q K L M A Q S G M S Q P D P S W P L L 300
919 TTTTTCCTCTTCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTGAGCCTACCTGAGCCCA 314
    F F P L L W A L P L L G P V *
979 TGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCTCTCTTGTGTTGGTAATAA
1039 ACACATTCCAGTTGATGCCCTTGCAGGGCATTCTCAAAAAAATAAAAAAATAAAAAA
1099 AAAAAAAAAAAAAAAAAA
```

SUBSTITUTE SHEET (RULE 26)

19/62

## Western blot of GST-Testisin using anti-Testisin peptide T175 antibody

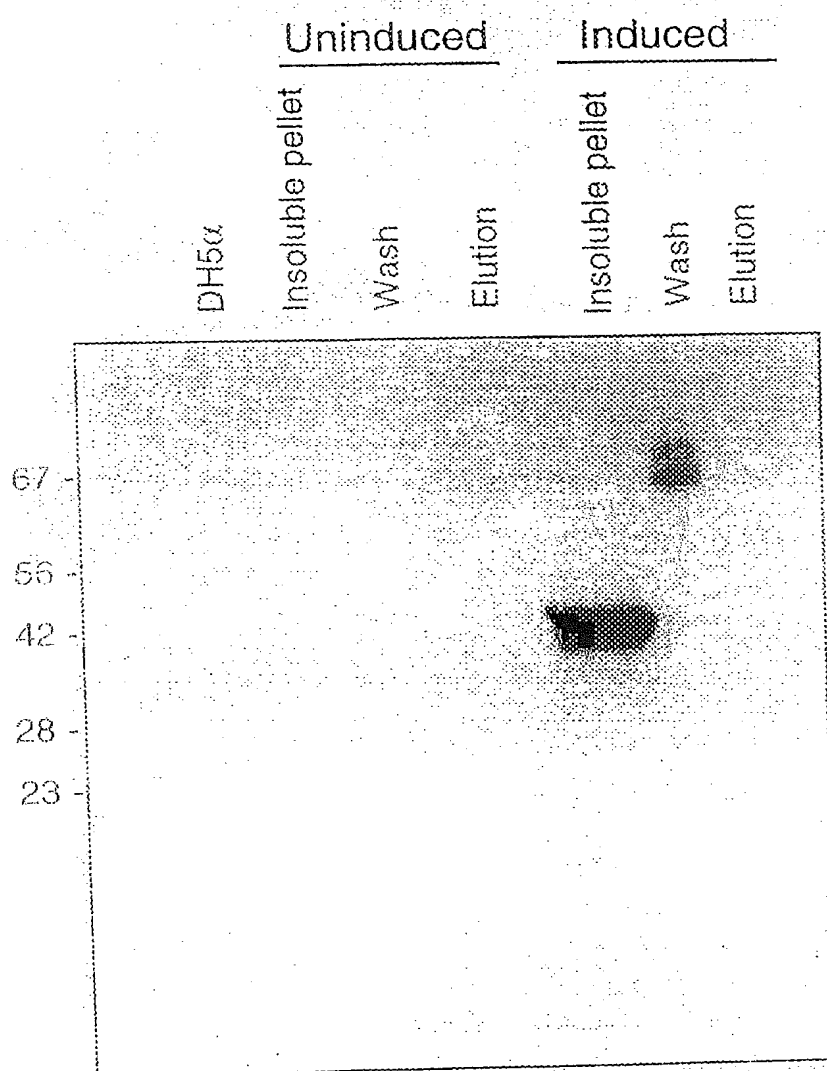


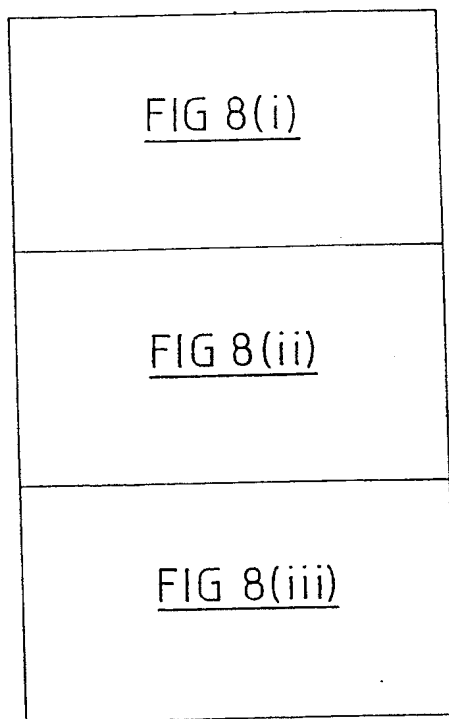
FIG 7

SUBSTITUTE SHEET (RULE 26)



20/62

FIG 8



21/62

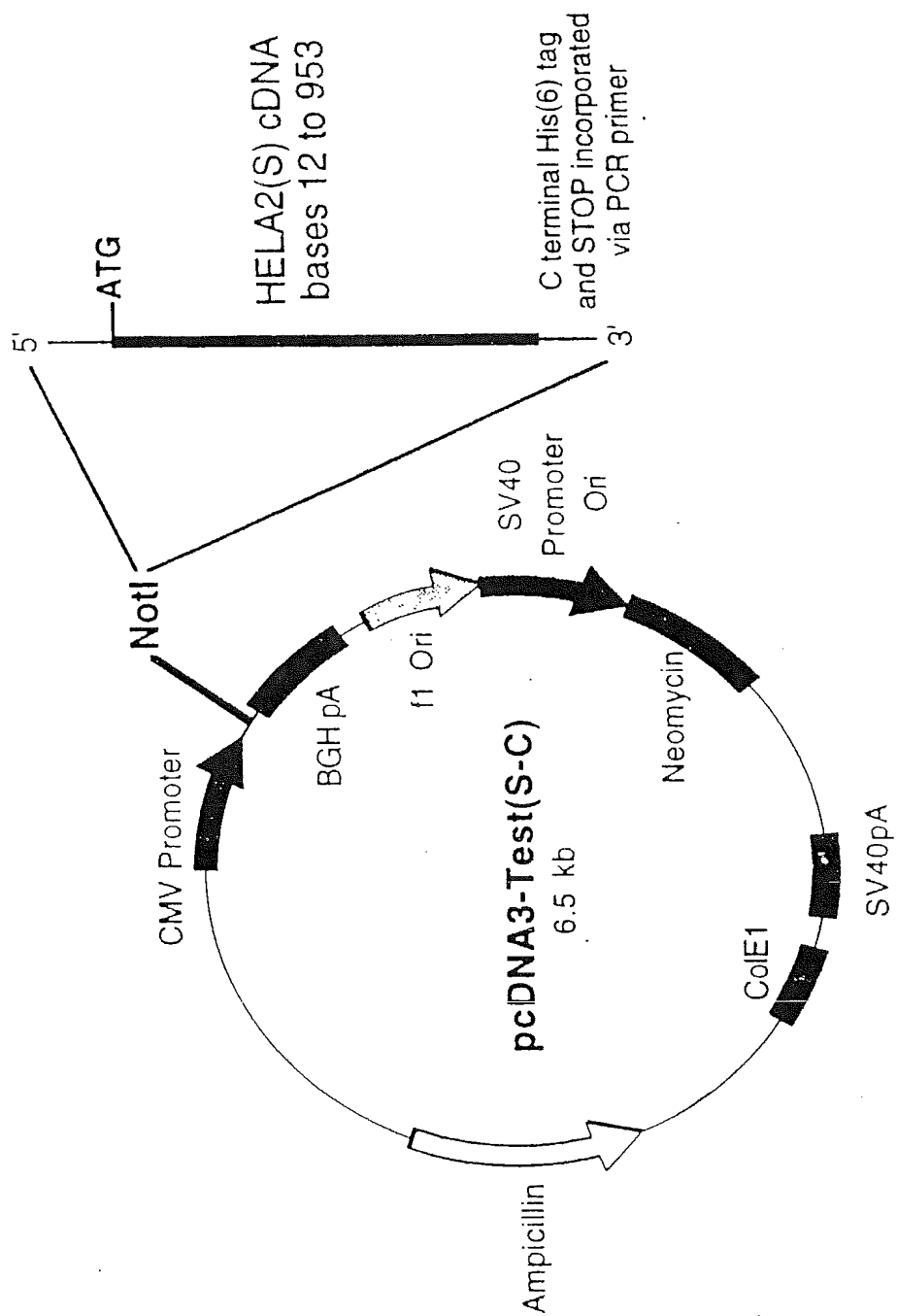
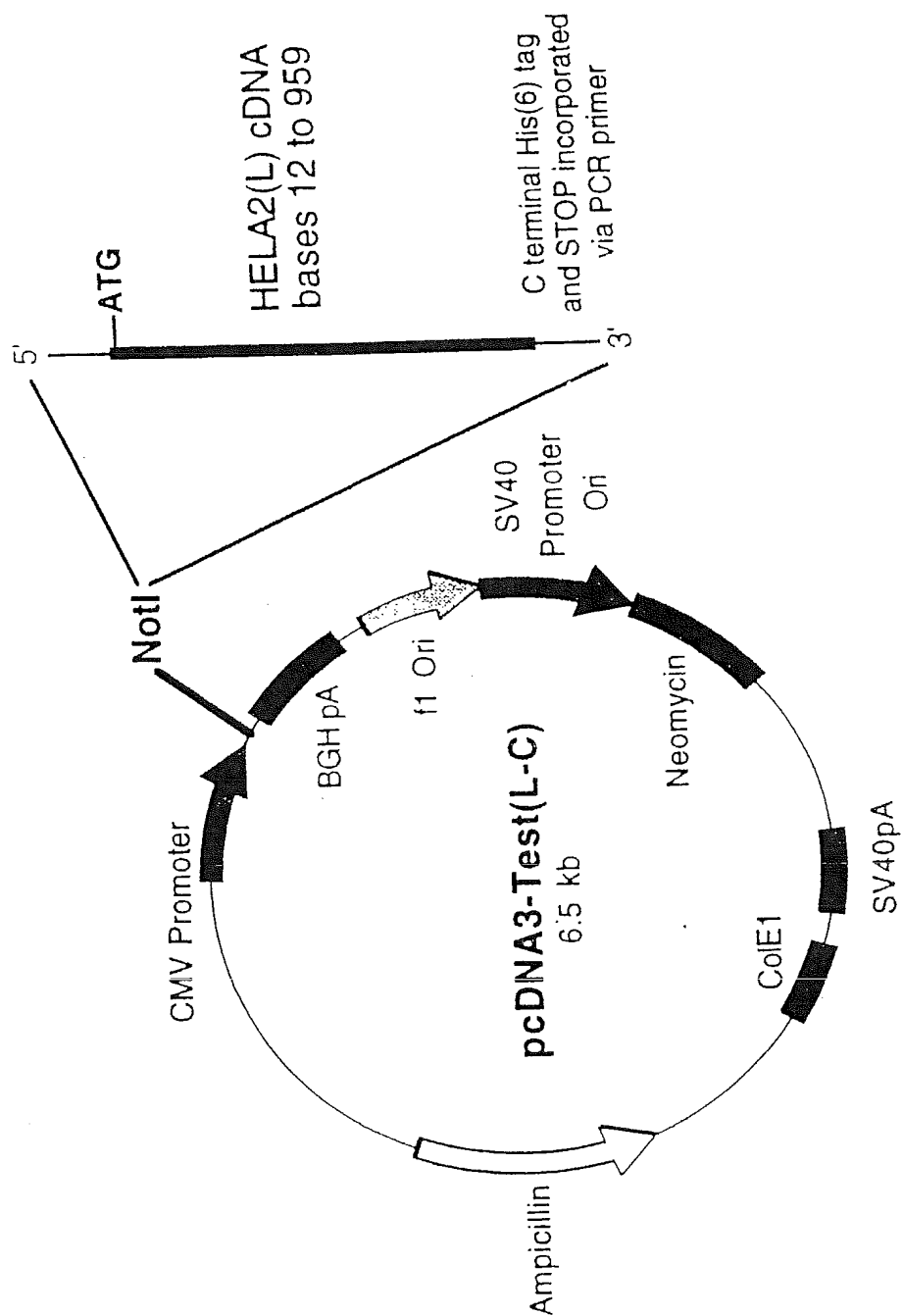


FIG 8(i)

SUBSTITUTE SHEET (RULE 26)

22/62

FIG 8(ii)

23/62

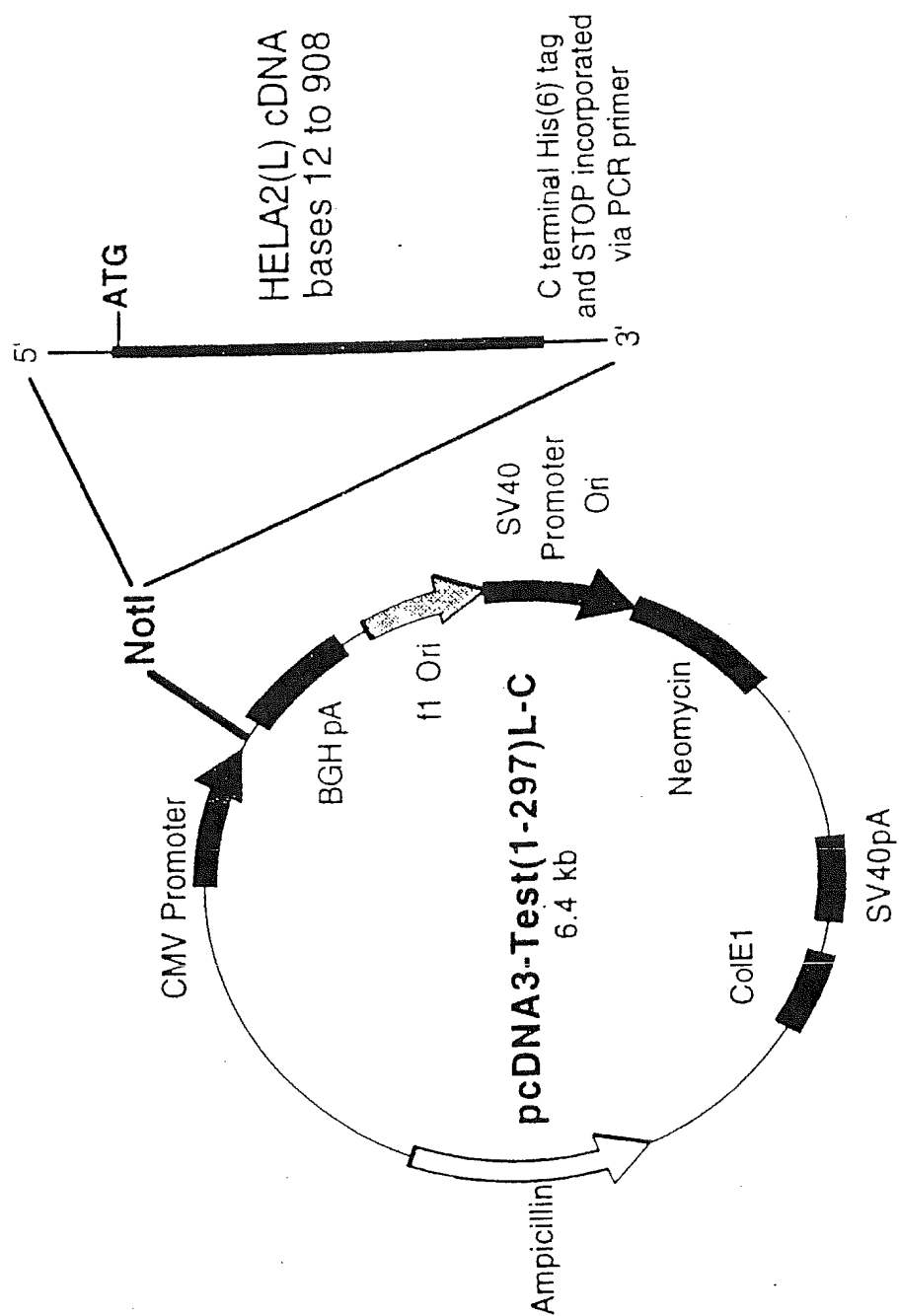


FIG 8(iii)

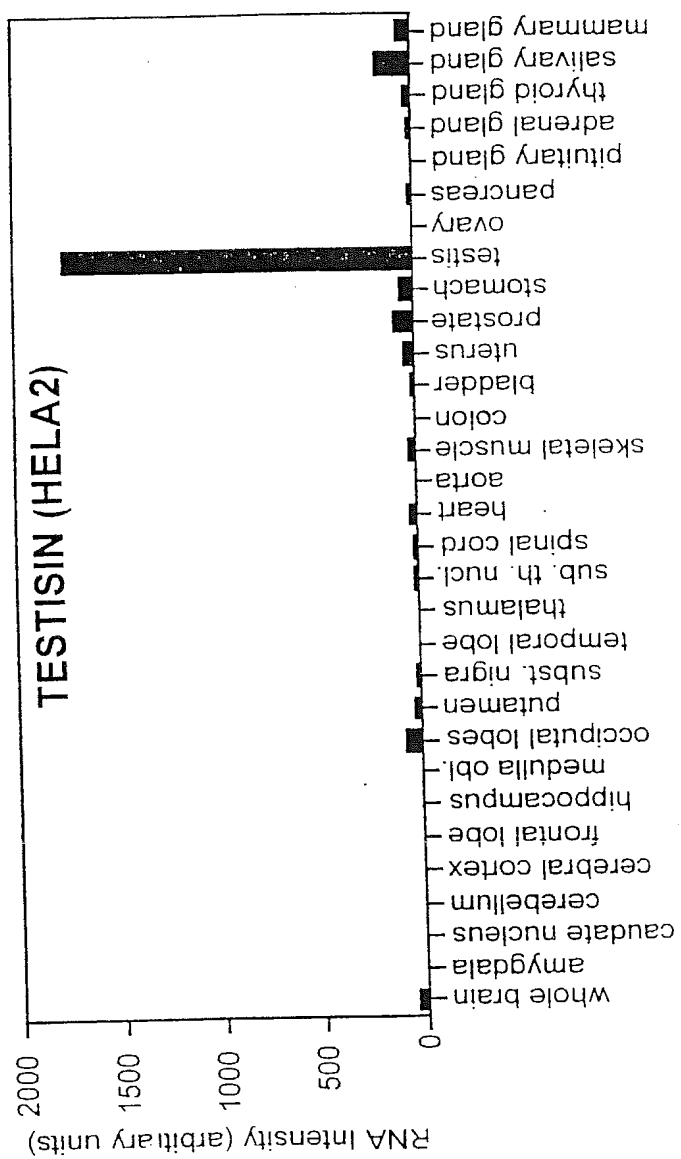
24/62

FIG 9

<u>FIG 9(i)</u>	<u>FIG 9(ii)</u>
<u>FIG 9(iii)</u>	<u>FIG 9(iv)</u>

25/62

FIG 9(i)



SUBSTITUTE SHEET (RULE 26)

26/62

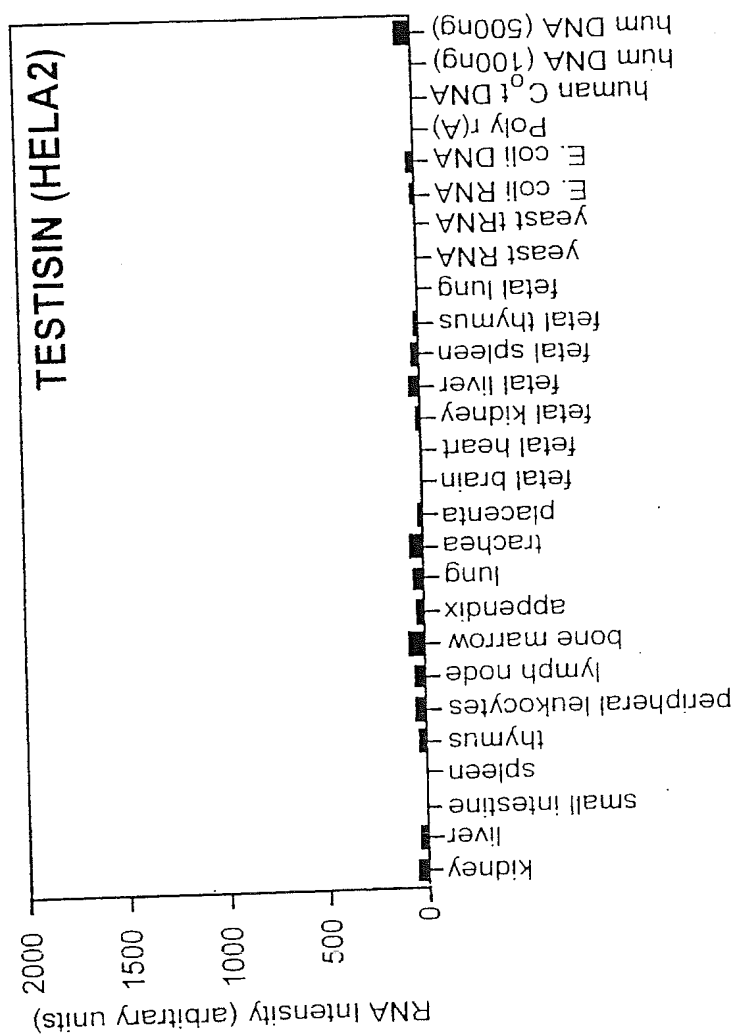
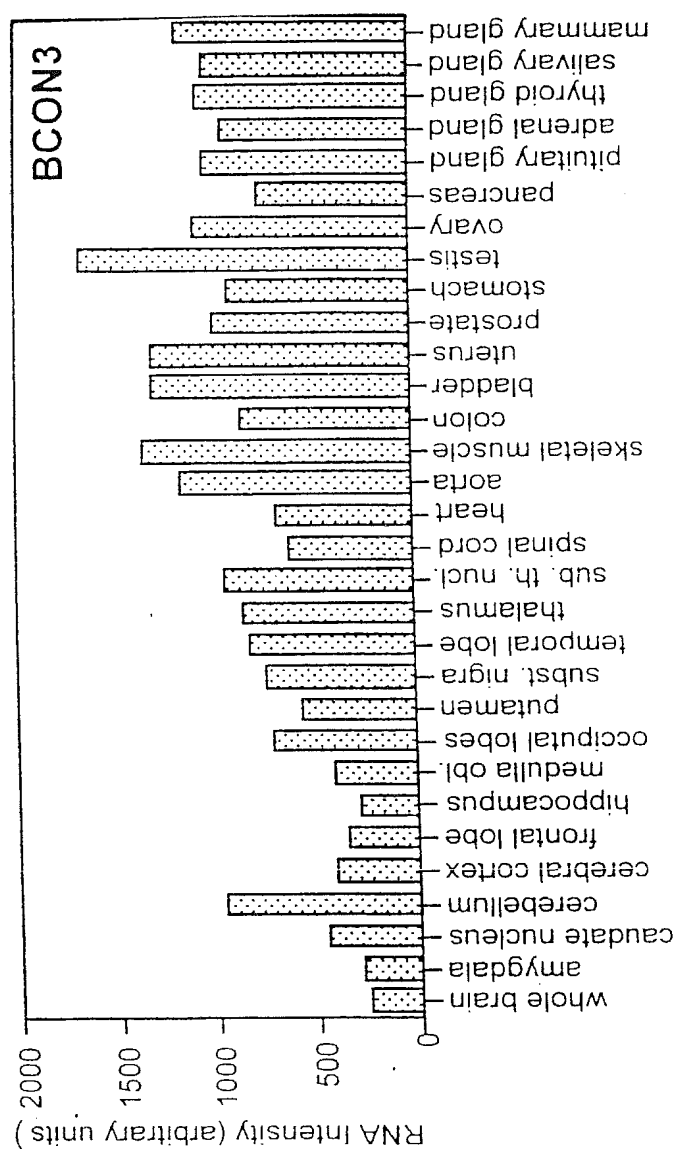


FIG 9(ii)

27/62

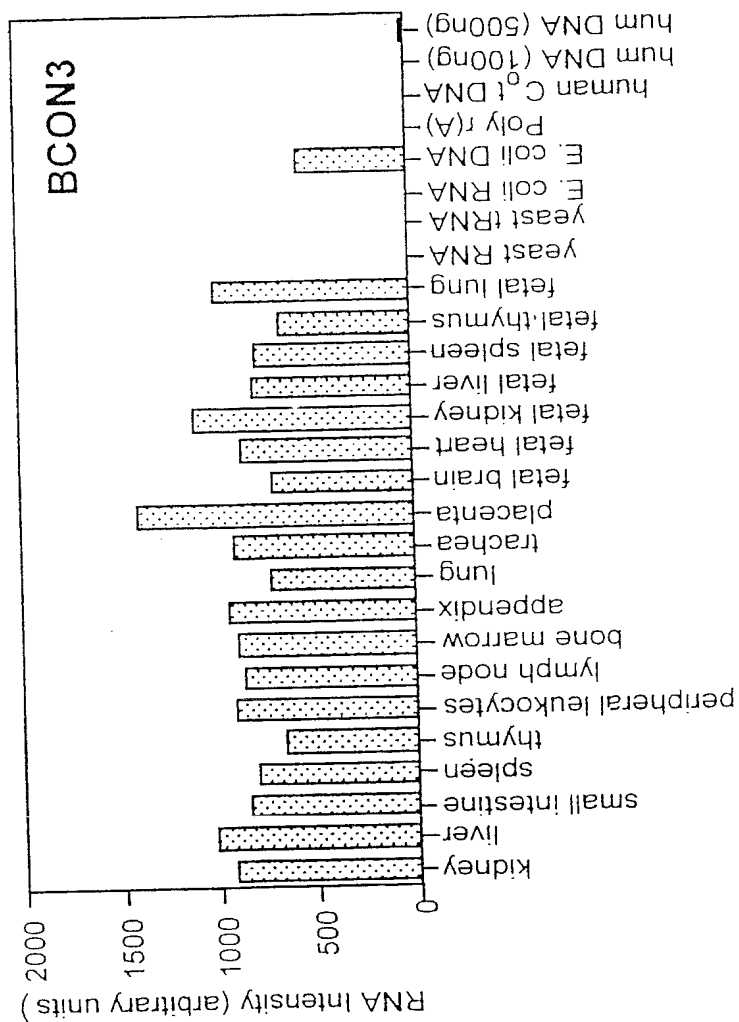
FIG 9(iii)



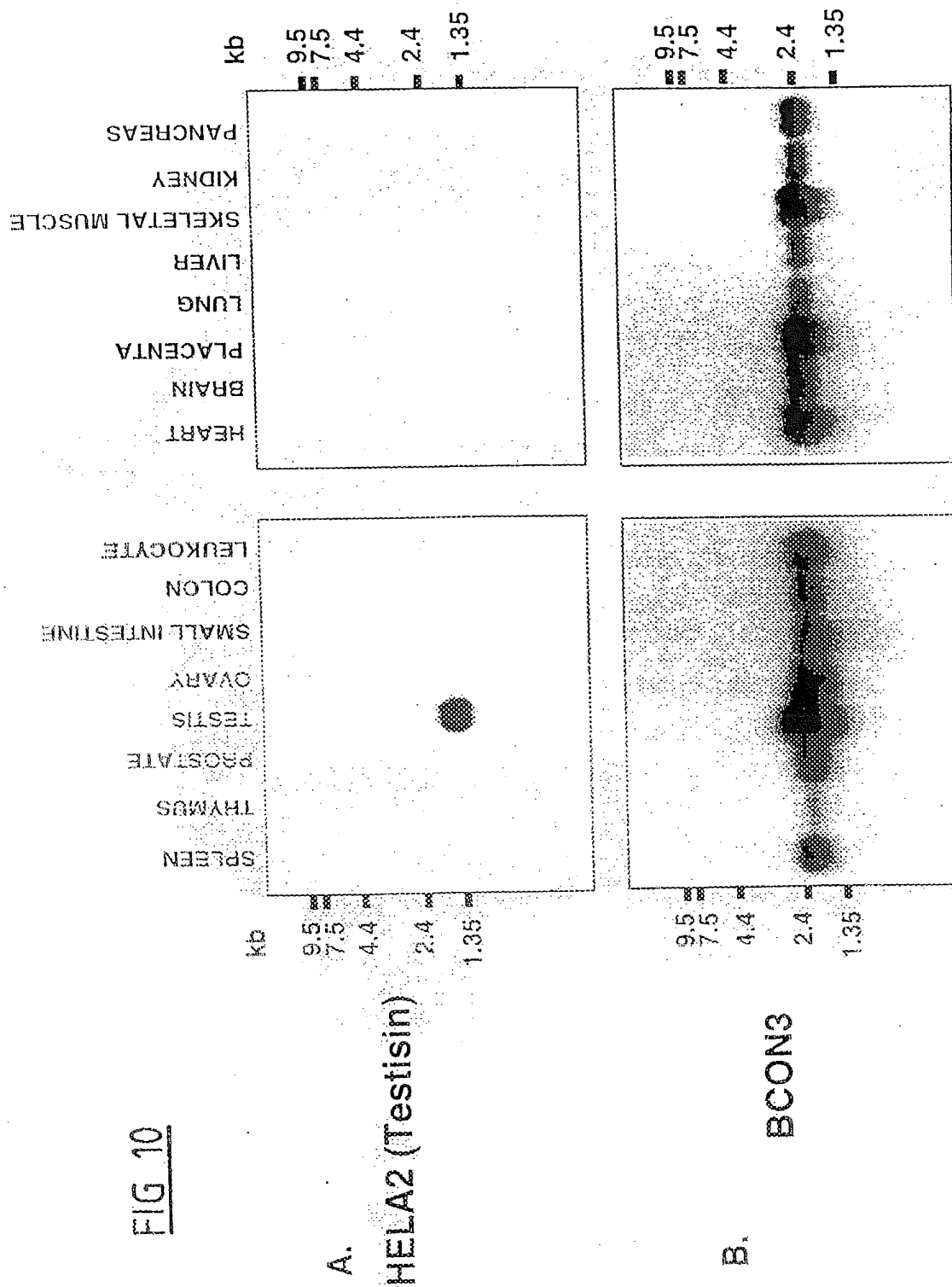


28/62

FIG 9(iv)

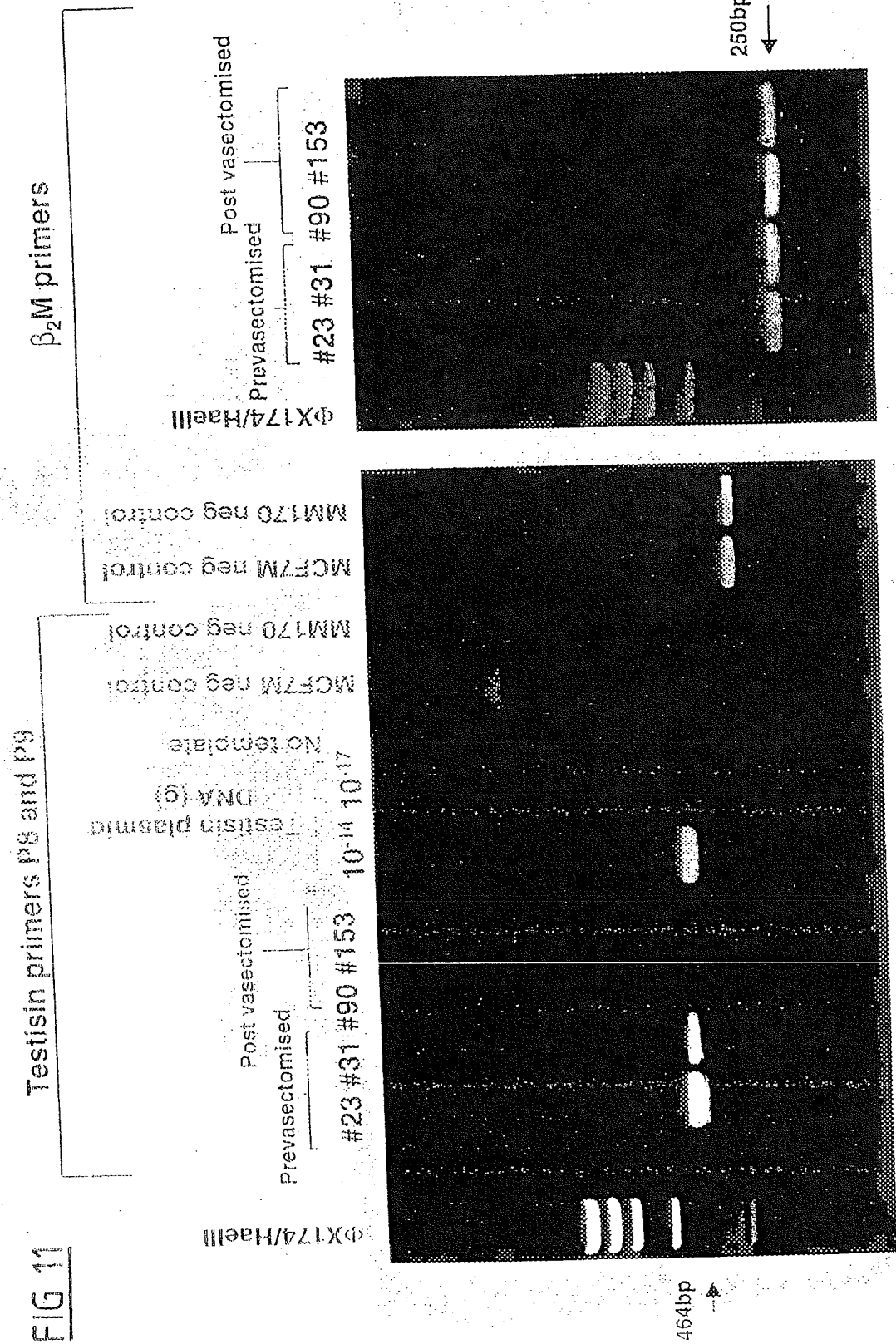


29/62



SUBSTITUTE SHEET (RULE 26)

30/62



31/62

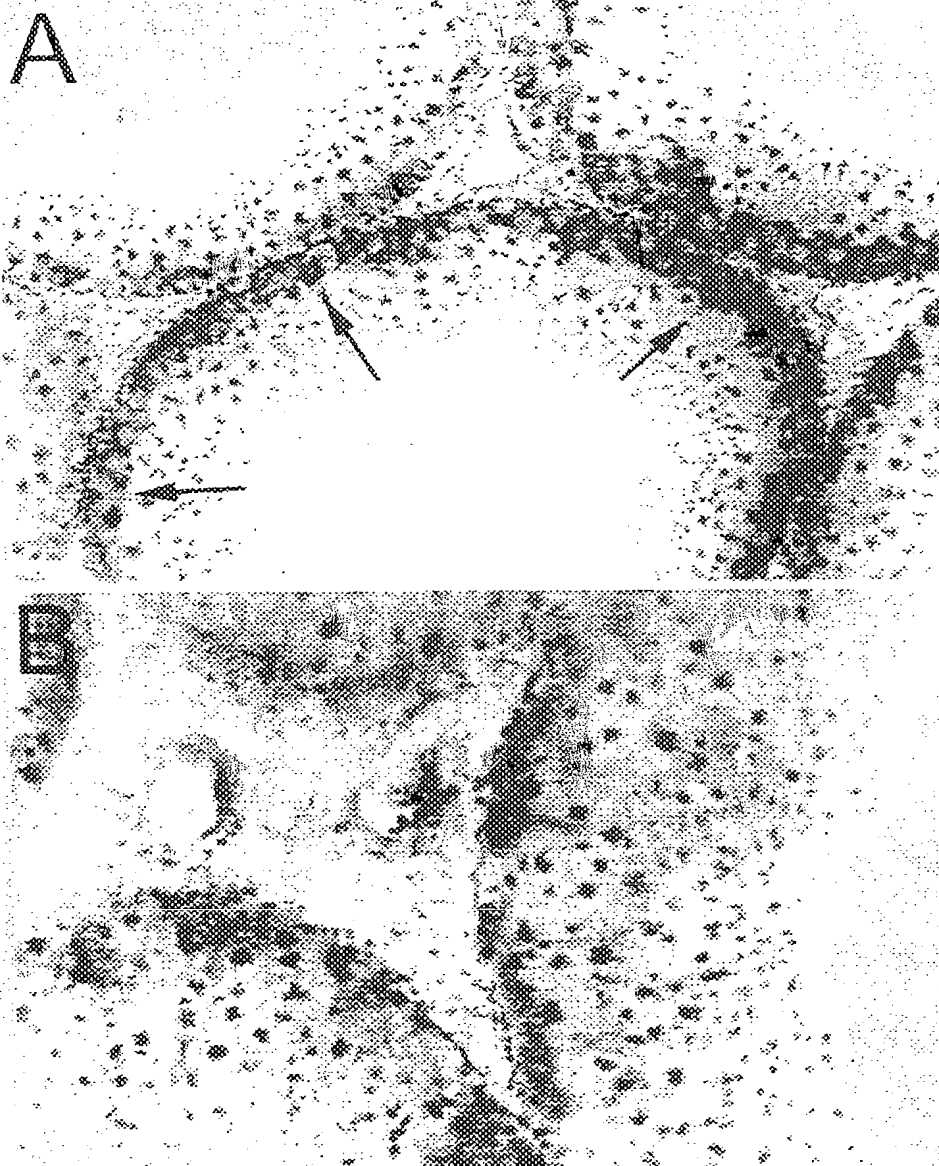


FIG 12

SUBSTITUTE SHEET (RULE 26)

32/62

Testis (HELA2) is located on human chromosome 16p13.3

A

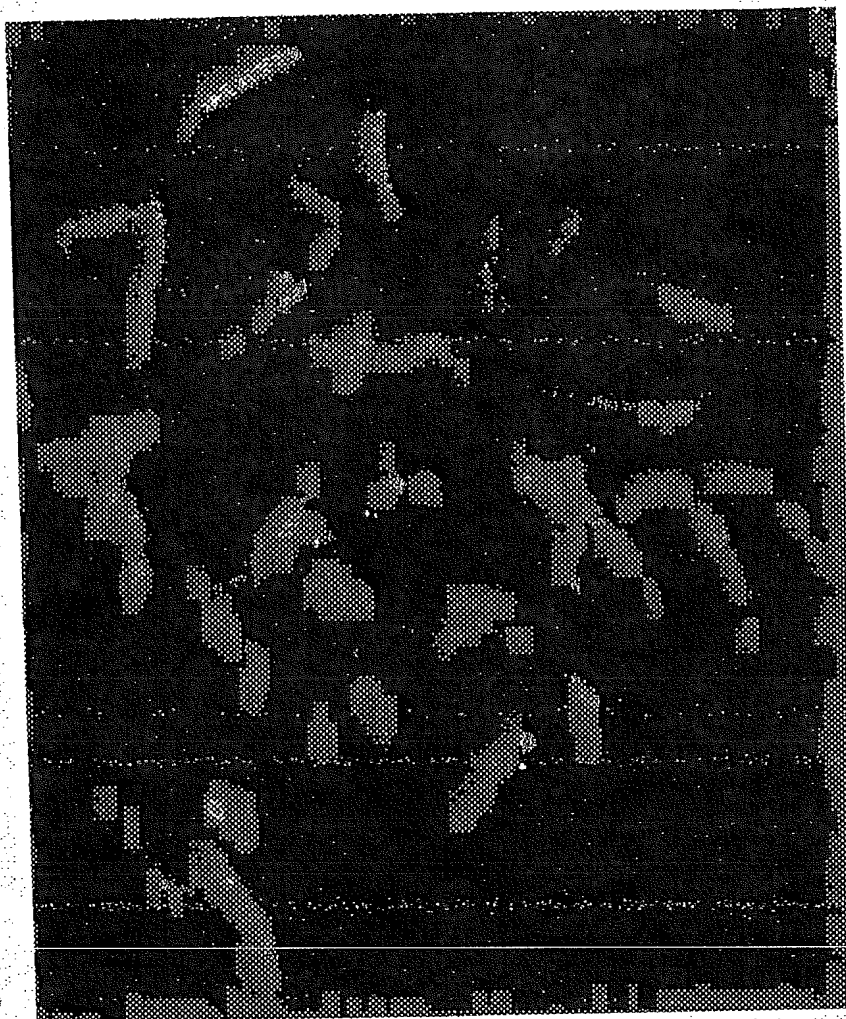
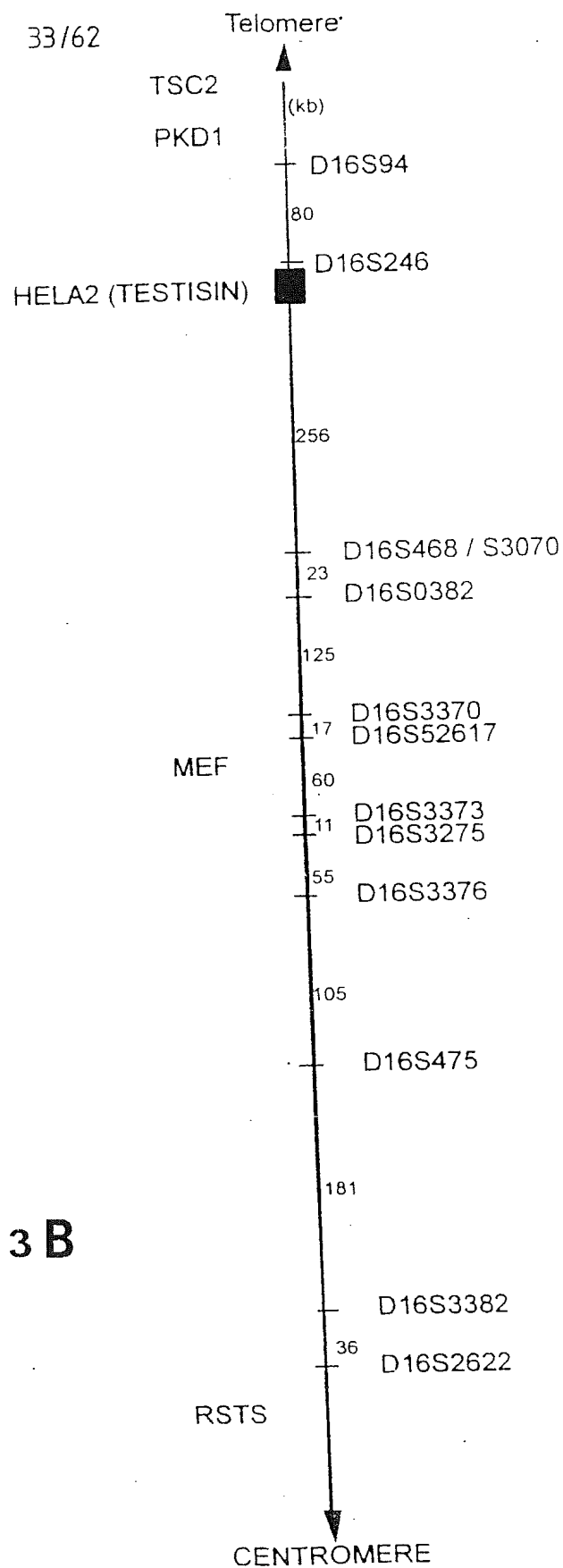


FIG 13A

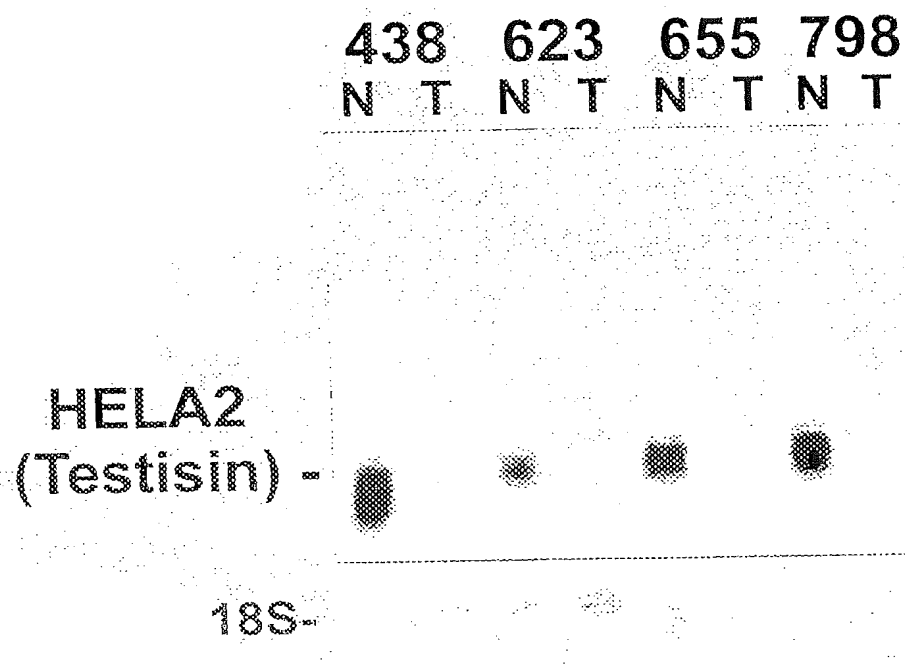
SUBSTITUTE SHEET (RULE 26)

**FIGURE 13 B**

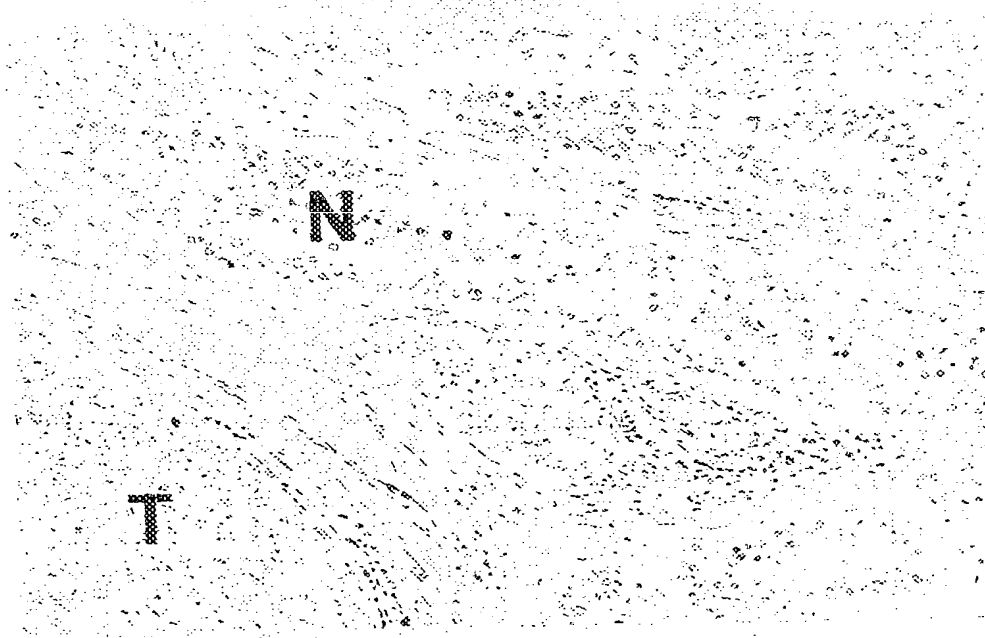
SUBSTITUTE SHEET (RULE 26)

FIG 14 34/62

## A. Northern Blot



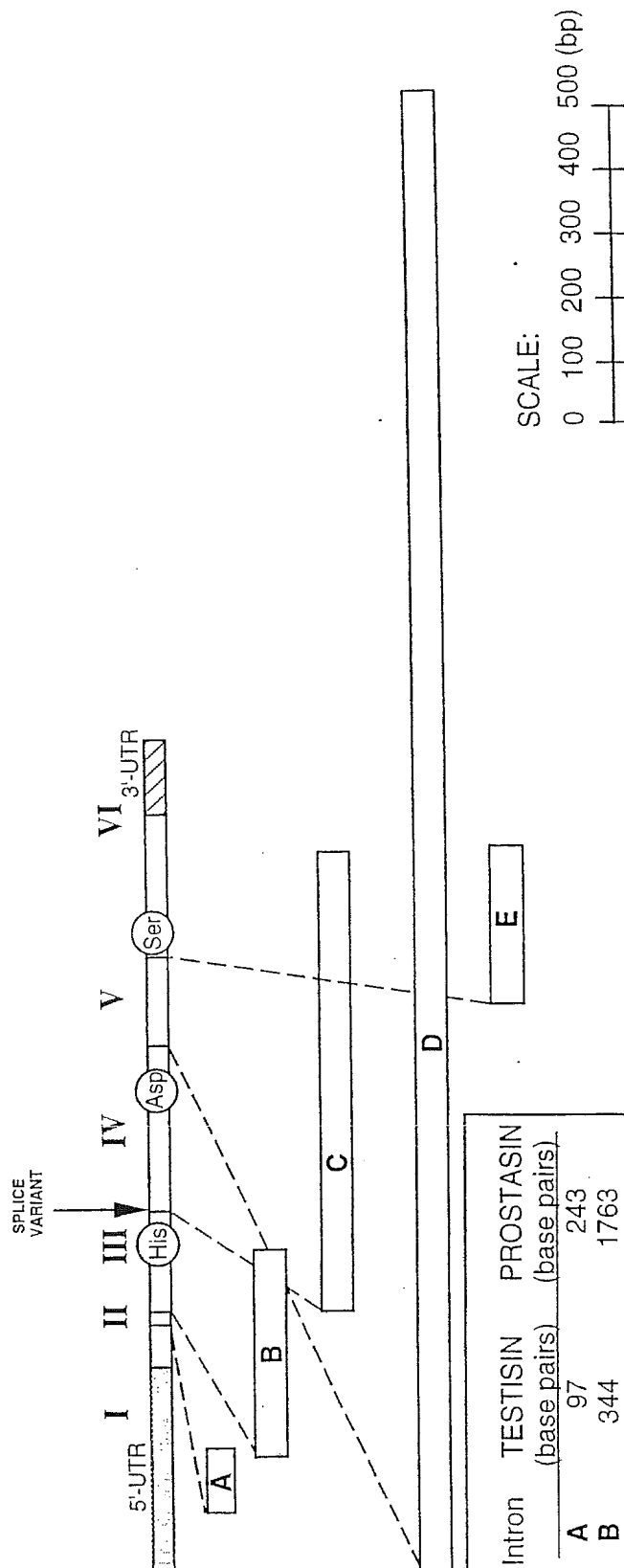
## B. Immunohistochemistry



SUBSTITUTE SHEET (RULE 26)

35/62

# TESTISIN INTRON/EXON BOUNDARIES AND SIZES



SCALE:

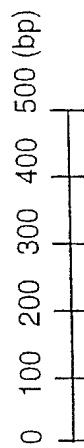


FIGURE 15

Intron	TESTISIN (base pairs)	PROSTASIN (base pairs)
A	97	243
B	344	1763
C	716	271
D	~2200	85
E	256	92

Exon	TESTISIN (base pairs)	PROSTASIN (base pairs)
I	>76	417
II	18	18
III	163	163
IV	284	272
V	168	167
VI	348	899



36/62

FIG 16

<u>FIG 16(i)</u>
<u>FIG 16(ii)</u>
<u>FIG 16(iii)</u>
<u>FIG 16(iv)</u>
<u>FIG 16(v)</u>
<u>FIG 16(vi)</u>

SUBSTITUTE SHEET (RULE 26)

37/62

```

50  agtgagtctc ctgcctcagc ctcccaagta gctgggactt caggtgtgtg
100 ccaccatcct cagctaattt tttttttttt tttttttttg agaaggagtc
150 ttgctctgtc gcccaggctg gagtgcagtg gcgcgatctt ccaggcccca
200 ccgggccctc aggaaggcct tgccctacctg ctttaagggg actcctggct
250 cagggccagg cccctggtgc tggaggaggt ggtgggtgga gggcaggggg
300 caccaagcgg gcagccagga cccccgggt gcagacaaga aaaggactgt
      /+1...EXON 1...
350 ggggtccacc ggtctgtggc cACATCAAGG AATGTGGTTG AAGACCCGCC
400 CTTAGGAGCT GAAAGCCAGG GCGCTACCAG GCCTGAGAGG CCCCAACAG
450 CCTTGGGCC TGGTTTGGGA GGATTAACT GGAGCTCCCA ACCCGCCCTG
500 CCCCCAGGG GCGACCCCGG GCCCGGCGG AGAGGAGGCA GAGGGGCGT
550 CAGGCCGCGG GAGAGGAGG CATGGGCGG CGCGGGCGG TGCTGCTGGC
      /INTRON A...
600 GCTGCTGCTG GCTCGGGCTG GACTCAGGAA GCCGGgtgag ctcggggcgc
650 tgctggcggg atggggaggc gggggagcgg tggggaggac gggagggtgga

```

FIG 16(i)

38/62

/EXON 2...		
ggccgcgggg	agtcacttct	tgtctccgcg agAGTCGCAG GAGCGGGCGC
/INTRON B...		
CGTTATCAG	tagggcgccc	aggacgcgcg attcctgccg gggccggttg
gccgaggtgg	acggggggcg	gtgagggggg agaggggggc cttactgct
ctctcgcccc	cgccccccgg	atcgagaact ctgttgcggt ggaaagtaac
taacggacgc	tggaggggga	tgggcgggcc ctgcagagca cgtgggagga
tctccagtgt	cacctacttc	ctgctgcaca cacgcgaggg gacctgggt
gggcaaaaac	gtgctttccc	ggacgggggtt gaaggggaga aaggagaggg
tcgggcttgg	ggggctgcct	cccgcggctc agcagttcct ctgaccatcc
/EXON 3...		
gagGACCATG	CGGCCGACGG	GTCATCACGT CGCGCATCGT GGGTGGAGAG
GACGCCGAAC	TCGGGCGTTG	GCCGTGGCAG GGGAGCCTGC GCCTGTGGGA
TTCCCCACGTA	TGCGGAGTGA	GCCTGCTCAG CCACCGCTGG GCACTCACGG

FIG 16(ii)

39/62

/INTRON C...  
CGGCGCACTG CTTTGAAACg tgagtggggg tgcgaacgga ggggtgcggg 1250  
gacgggcagg aacagggctg gagggagtgc caccgaactt tacctctggt 1300  
ctgatgccag acttgggcgt gaaagtgtg cgtggatgcg gcctgggtgtt 1350  
ctcctgagcc ccaggctgtg ctgcagccgg ttacaccac tccagttccc 1400  
tttgggtctc ctggaggga ccctgttcag gttattccag aatgttcttc 1450  
cagaacattt ccacacatt ttgggtattc tctccctttt tctttcaacc 1500  
caaagttcac cactgaccat cccaccctca tccccctcc tggtaggacgg 1550  
tgcggtacag tgtggggcac tgagccaaagg ccagcaccac cgggccgctg 1600  
tgtggactcc atcctgccaa tcccacattg gcgtgggtgca tctcccatt 1650  
cctccttggg ctgcatgggg gtgccccctgg aggccttggc tcaatgcaag 1700  
gctccttggg acagctctgg gaggtgacaa gacccaccc ttctgctgca 1750  
ggagcaggtc ctaggacttt ggttgtggtc tgtctgggct ccttcatttc 1800  
tgcaggggac cctgggtgtt agcaagtagc agcaacacca cagtttcccc 1850  
tcctgcactg gacccagtt gtgctcaggt agccagccct ccatccaggg 1900

FIG 16(iii)

40/62

## /EXON 4...

ccccgactg ctctcttctc ttctgccagc tatagTGACC TTAGTGATCC 1950  
CTCCGGGTGG ATGGTCCAGT TTGGCCAGCT GACTTCCAAG CCATCCTTCT 2000  
GGAGCCTGCA GGCCTACTAC ACCCGTTACT TCGTATCGAA TATCTATCTG 2050  
AGCCCTCGCT ACCTGGGGAA TTCACCCCTAT GACATTGCCT TGGTGAAGCT 2100  
GTCTGCACCT GTCACCTACA CTAACACAT CCAGCCCATC TGCTCCTCAGG 2150  
CCTCCACATT TGAGTTTGAG AACCGGACAG ACTGCTGGGT GACTGGCTGG 2200

## /INTRON D...

GGGTACATCA AAGAGGATGA GGgtgaggct ggggacaggc ggggtcagggg 2250  
ggaactgtct ttgttcacct gtccccctgc ataggcacia tagccccctg 2300  
cttggtcttg ggggtgcaggc tatgccccctc ttgcttgagc tctctcctca 2350  
cctgccaggg cagggaccaa acacccagtt ctctcccttc caggggctgt 2400  
ggggggccaga aggagagtgt gagaggagg ccagtttggc gcaagcctgt 2450  
gggtggtgcg gtggtggagg ggcttctggag ggcttggcga cataaacctc 2500  
atacttggat ttattcctgc atctttccac ctccccagc gctcaccaat 2550

FIG 16(iv)

41/62

```

gccccaggca tca.....approx 1000 bp.....
ccagggtgcc cttcccccaggcttggct ttggatgctt atgtgaacac
cgttttaagt tgccttgcc ccttcctcgg ttcctttttg gctgaggaat
ctctccatgg ctgcaggcag ggcattgtt gccattctac agatagggaa
agtgcggctg ggggagctct gacagctgtc cctccccggg gccttctgtg
atgtctgctga gggcctctgt tgtgctgggg tctgggttgg agctgggggt
aatggagatg aacctgccag gcacagtggg tgccccaggg cccccacccc
cgcagcctat gccatccctc catagagggg cctcaggttg ctgtctctct
                                /EXON 5...

ccttcccact atcgtccgca cagCACTGCC ATCTCCCCAC ACCCTCCAGG
AAGTTCAGGT CGCCATCATA AACAACTCTA TGTGCAACCA CCTCTTCCTC
AAGTACAGTT TCCGCAAGGA CATCTTTGGA GACATGGTTT GTGCTGGCAA
                                /INTRON E...

TGCCCCAAGGC GGAAGGATG CCTGCTTCgt gagtgctcctt gccaccactc
ccagccccagg aaagcctcct gtgtccctgt gccttatttg accctcatgc
caacccccggg aggtggagac tgttgcccc ctctgcagat gcagaaacgg

```

FIG 16(v)

42/62

```

aggcttggct gctgccaggg ggaggaggag gatgtgcacc cagtctaccc ~4263
agccccatag cccttcccac tctcagcccc tccccgtccc cactcactct ~4313
                               /EXON 6...

gccccaggct gacctcagcc ccgctgctcc ccagGGTGAC TCAGGTGGAC ~4363
CCTTGGCCTG TAACAAGAAT GGACTGTGGT ATCAGATTGG AGTCGTGAGC ~4413
TGGGGAGTGG GCTGTGGTCG GCCCAATCGG CCCGGTGTCT ACACCAATAT ~4463
CAGCCACCAC TTTGAGTGGA TCCAGAAGCT GATGGCCCAG AGTGGCATGT ~4513
CCCAGCCAGA CCCCTCCTGG CCGCTACTCT TTTTCCCTCT TCTCTGGGCT ~4563
CTCCCACCTCC TGGGGCCGGT CTGAGCCCTAC CTGAGCCCAT GCAGCCTGGG ~4613
GCCACTGCCA AGTCAGGCCC TGGTTCTCTT CTGTCTTGTT TGGTAATAAA ~4663
CACATTCCAG TTGATGCCCTT GCAGGGCATT CTTCaaaagc agtgggcttca ~4713
tggacagctc attctctctt gtgcagacag cctgtctgtg cccctggctc ~4763
acacccacat ctgttctgca ccatagaacc atctggttat ttcgatcaga ~4813
aagagaattg tgtgttgccc aggcctgggtct tgaacgccta ggggtgtctcg ~4863
atc

```

FIG 16(vi)

43/62

EXON III CACTGCTTTGAAAC**gt**gagtgggggtgcgaacggag  
 ggggtgcggggacggggcaggaacaggggctggagggagtgccaccga  
 actttacctctggtctgatgccagacttgggcgtgaaagttgtgc  
 gtggatgcggcctggtgttctcctgagccccaggctgtgctgcag  
 ccggttacaccactccagttccctttgggtctcctggaggggaac  
 cctgttcagggttattccagaatgttcttccagaacatttccacac  
 acttttgggtattctctccctttttcttttcaacccaaagttcacc  
 actgaccatcccaccctcatccccctcctgggtggacgggtgcggt  
 acagtgtggggcactgagccaaggccagcacccccgggcccgtgt

.....INTRON C (716 BP).....

gtggactccatcctgccaatcccacattggcgtggtgcatctccc  
 cattcctccttgggctgcatgggggtgcccctggagggccttggct  
 caatgcaaggctccttgggacagctctgggaggtgacaagacccc  
 acccttctgctgcaggagcaggctcctagactttggttgtggtctg  
 tctgggctccttcatttctgcaggggaccctgggtgttagcaagt  
 agcagcaacaccacagtttcccctcctgcactggaccccagttgt  
 gctcaggtagccagccctccatccagggcccctgactgctctctt  
 ctcttctgccc**ag**ctat**ag**TGACCTTAGTGATCCC EXON IV

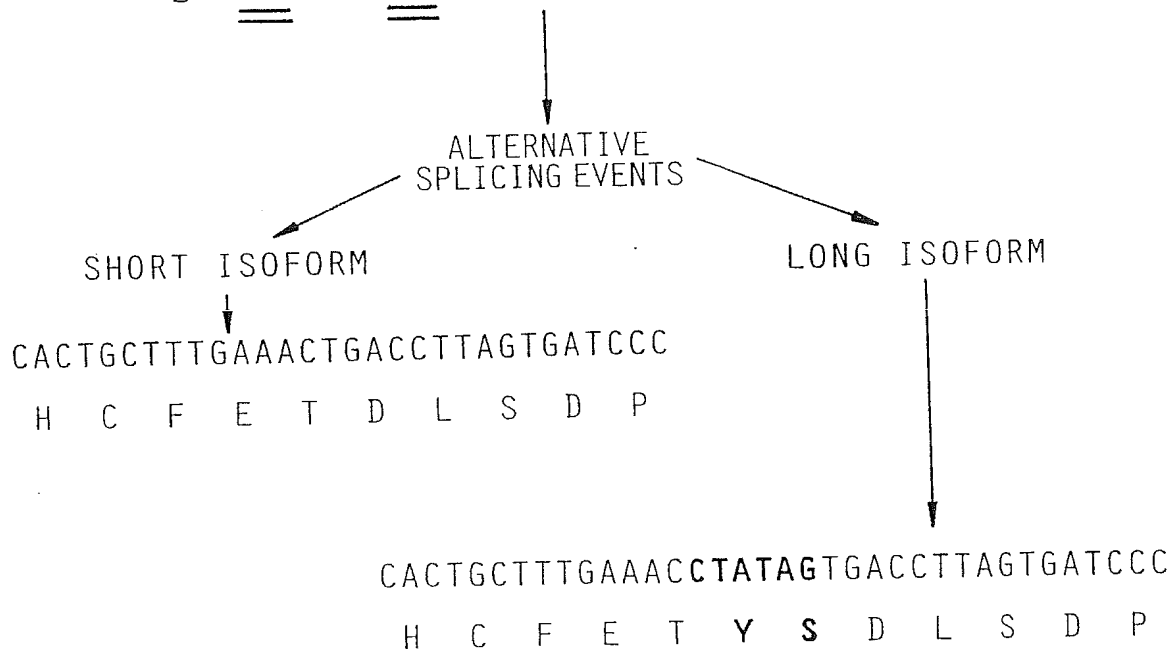


FIGURE 17



44/62

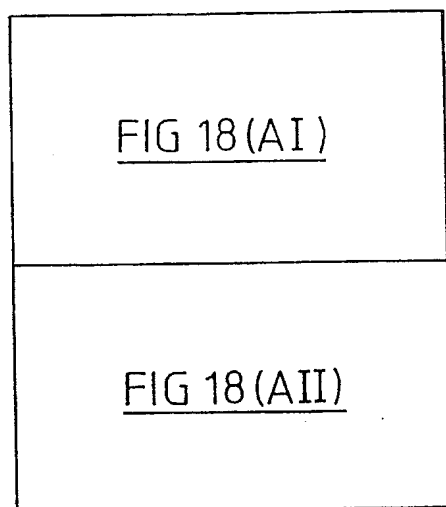


FIG 18(A)

SUBSTITUTE SHEET (RULE 26)

45/62

## FIGURE 18 (AI)

1 CGACCTATTGTCAGGGCCCTGCGGTACAGGACCATCCCTTCCCGTATAGTGGTGGCGG 20  
D L L S G P C G H R T I P S R I V G G D  
61 TGATGCTGAGCTTGGCCCGCTGGCCCGTGGCAAGGAGCCCTGCGTGTATGGGCAACCACCTT 40  
D A E L G R W P W Q G S L R V W G N H L  
121 ATGTGGCGCAACCTTGTCTCAACCGCCGCTGGGTGCTTACAGCTGCCCACTGCTTCCAAAA 60  
C G A T L L N R R W V L T A A H C F Q K  
181 GGATAACGATCCCTTTTGACTGGACAGTCCAGTTTGGTGAGCTGACTTCCAGGCCATCTCT 80  
D N D P F D W T V Q F G E L T S R P S L  
241 CTGGAACCTACAGGCCCTATTCCAACCGTTACCAAATAGAAGATATTTCTGAGCCCCAA 100  
W N L Q A Y S N R Y Q I E D I F L S P K  
301 GTACTCGGAGCAGTATCCCAATGACATAGCCCTGCTGAAGCTGTCACTCTCCAGTCACCTA 120  
Y S E Q Y P N D I A L L K L S S P V T Y  
361 CAATAACTTCATCCAGCCCATCTGCCCTCCTGAACCTCCACGTACAAGTTTGAGAAACCGAAC 140  
N N F I Q P I C L L N S T Y K F E N R T  
421 TGA CTGCTGGGTGACCGGCTGGGGGGCTATTGGAGAAGATGAGAGTCTGCCATCTCCCAA 160  
D C W V T G W G A I G E D E S L P S P N

46/62

**FIGURE 18 (AII)**

481 CACTCTCCAGGAAGTGCAGGTAGCTATTATCAACAACAGCATGTGTAACCATATGTACAA  
T L Q E V Q V A I I N N S M C N H M Y K 180

541 AAAGCCAGACTTCCGCACGAAACATCTGGGAGACATGGTTTGGCGTGGCACTCCTGAAGG  
K P D F R T N I W G D M V C A G T P E G 200

601 TGGCAAGGATGCCCTTGGTGACTCGGAGGACCCCTTGGCCCTGCGACCGAGGATACGGT  
G K D A C F G D S G G P L A C D Q D T V 220

661 GTGGTATCAGGTTGGAGTTGTGAGCTGGGGAATAGGCTGTGTCGCCCAATCGCCCTGG  
W Y Q V G V S W G I G C G R P N R P G 240

721 AGTCTATACCAACATCAGTCATCACTACTAACAAGTCCAGTCAACCATGATCCGCAATGG  
V Y T N I S H H Y N W I Q S T M I R N G 260

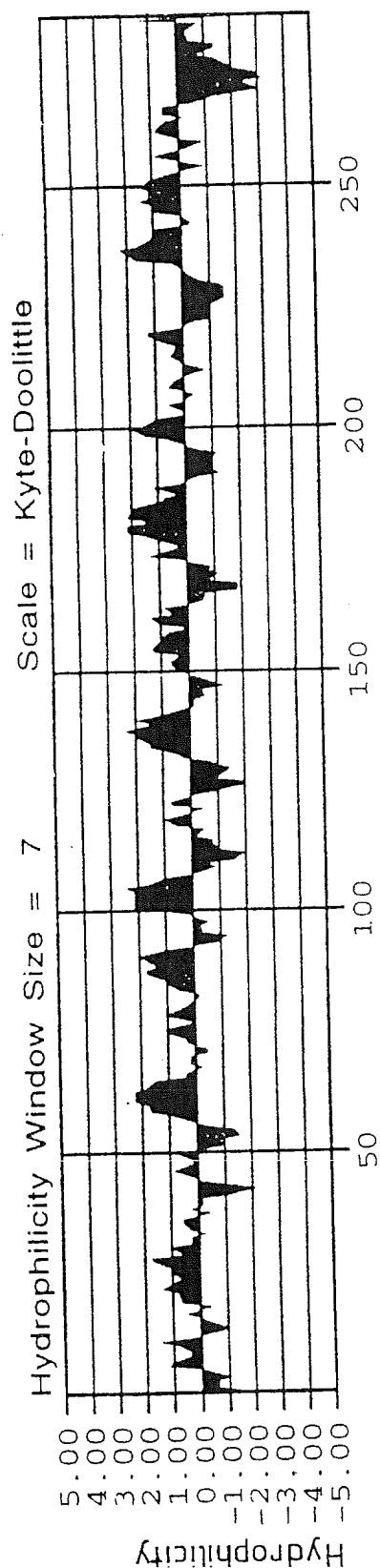
781 GCTGCTCAGGCCTGACCCAGTCCCCTTGCTACTGTTTCTTACTCTGGCCCTGGGCTTCCTC  
L L R P D P V P L L L F L T L A W A S S 280

841 TTTGCTGAGGCCTGCCCTGAGCCACACGTTACGTACACACCTGTGAGGTCAGGGTGTGTC  
L L R P A 285

901 TCCTTTGTATCTTGCTTGCTAATAAACCTGTTAATATTAAAAAATAAAAAA

47/62

FIG 18B



SUBSTITUTE SHEET (RULE 26)



49/62

FIG 20A(AI)

FIG 20A(AII)

FIG 20A(AIII)

FIG 20A(A)

SUBSTITUTE SHEET (RULE 26)

50/62

## FIGURE 20A (AI)

1 CTGAACCGGTTGTGGCGCGGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGAGC 60  
L N R  $\nabla$  V V G G E D S T D S E W P W I V S

21 ATCCAGAAGAAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC 120  
I Q K N G T H H  $\square$  A G S L L T S R W V I

41 ACTGCTGCCCACTGTTTCAAGGACAAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTG 180  
T A A  $\textcircled{H}$   $\square$  F K D N L N K P Y L F S V L L

61 GGGCCCTGGCAGCTGGGGAACCCCTGGCTCTCGGTCCAGAAAGTGGGTGTTGCCCTGGGTG 240  
G A W Q L G N P G S R S Q K V G V A W V

81 GAGCCCAACCCCTGTGTATTCTGGAAGGAAGGTGCCCTGTGCAGACATTGCCCTGGTGCGT 300  
E P H P V Y S W K E G A C A  $\textcircled{D}$  I A L V R

101 CTCGAGCGCTCCATACAGTTCTCAGAGCGGGTCCCTGCCCATCTGCCCTACCTGATGCCCTCT 360  
L E R S I Q F S E R V L P I  $\square$  L P D A S

121 ATCCACCTCCCTCCAAACACCCCACTGCTGGATCTCAGGCTGGGGAGCATCCAAGATGGA 420  
I H L P P N T H  $\square$  W I S G W G S I Q D G

SUBSTITUTE SHEET (RULE 26)

51/62

## FIGURE 20A (AII)

GTTCCCTTGCCCCACCCCTCAGACCCCTGCAGAAGCTGAAGTTCTCTATCATCGACTCGGAA 480  
 141 V P L P H P Q T L Q K L K V P I I D S E  
  
 GTCTGCAGCCATCTGTACTGGGGGAGCAGACAGGACCCATCACTAGGACATGCTG 540  
 161 V [C] S H L Y W R G A G Q G P I T E D M L  
  
 TGTGCCGGCTAACTTGGAGGGGAGCGGGATGCTTGTCTGGCGCACTCCGGGGCCCCCTC 600  
 181 [C] A G Y L E G E R D A [C] L G D [S] G G P L  
  
 ATGTGCCAGGTGGACGGCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGCTGT 660  
 201 M [C] Q V D G A W L L A G I I S W G E G [C]  
  
 GCCGAGCGCAACAGCCCGGGTCTACATCAGCCCTCTCTGCGCACCGCTCCTGGGTGGAG 720  
 221 A E R N R P G V Y I S L S A [H] R S W V E  
  
 AAGATCGTGCAAGGGGTGCAGCTCCGCGGGCGGCTCAGGGGGGTGGGGCCCTCAGGGCA 780  
 241 K I V Q G V Q L R G R A Q G G A L R A  
  
 CCGAGCCAGGCTCTGGGGCCGCGGCTCCTAGGGCCCAAGCGGACGCGGGGCTCGG 840  
 261 P S Q G S G A A R S  
  
 ATCTGAAAGCGGCAGATCCACATCTGGATCTGGATCTGCGGGCGCCCTCGGGCGTTTC 900  
 CCGCGCGTAAATAGGCTCATCTACCTCTACCTCTGGGGGCGCGGACGGCTGCTCGGAA 960

SUBSTITUTE SHEET (RULE 26)



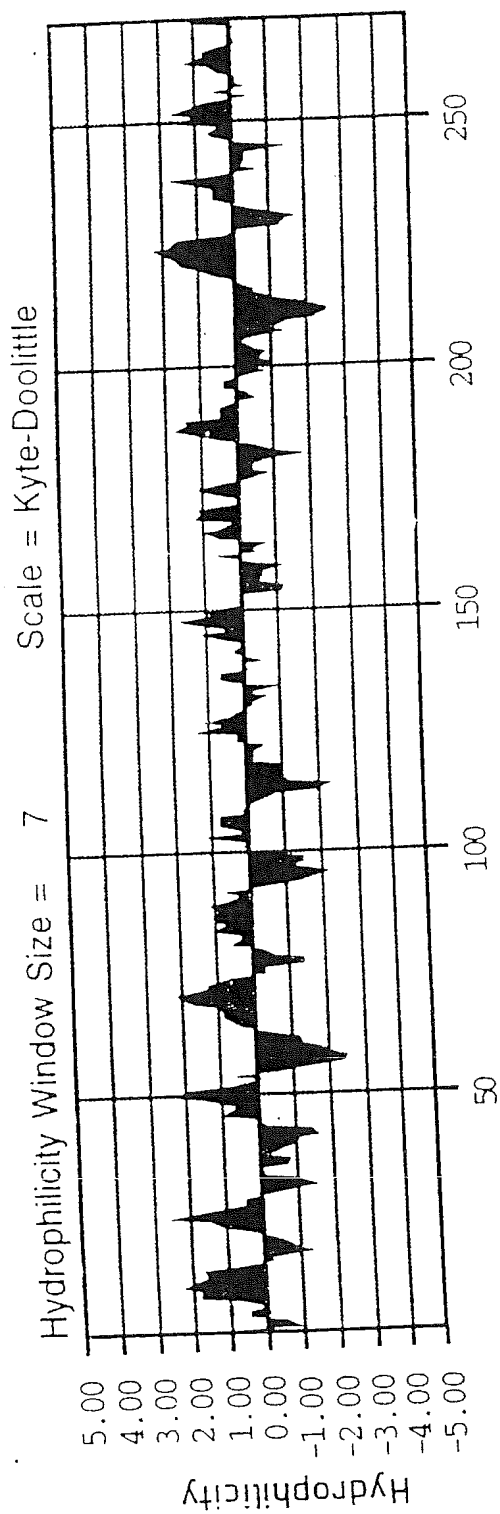
52/62

**FIGURE 20A (AIII)**

AGGAAACCCCTCCCGACCCGACGGCCCTCAGGCCCGCCCTCCAAGGCATCAGGCC 1020  
CCGCCCAACGGCCTCATGTCCCGCCCCACGACTTCCGGCCCCCGGGCCCGCAGCG 1080  
CTTTGTGTATATAAATGTTAATGATTTTATAGGTATTGTAAACCCCTGCCACATATCT 1140  
TATTATTCCCTCCAATTCAATAA

53/62

FIG 20A (B)



54/62

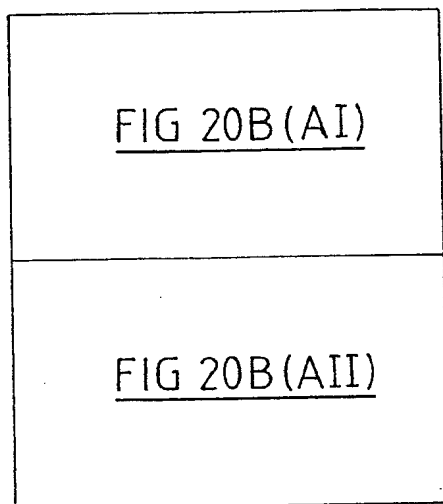


FIG 20B(A)

SUBSTITUTE SHEET (RULE 26)

55/62

## FIGURE 20B (AI)

1 AATGCGGCCACTCCAAGGAGCGCGGAGGATTGTGGAGGCCAAGACACCCAGGAAGAC 60  
[C] G H S K E A G R V I V G G Q D T Q E G

21 R W P W Q V G L W L T S V G H V [C] G G S  
GCTGGCCGTGGCAGGTTGGCCGTGGTTGACCTCAGTGGGGCATGTATGTGGGGGCTCCC 120

41 L I H P R W V L T A A (H) [C] F L R S E D P  
TCATCCACCCACGCTGGGTGCTCACAGCCGCCCACTGCTTCTGAGGTCTGAGGATCCCG 180

61 G L Y H V K V G G L T P S L S E P H S A  
GGCTCTACCATGTTAAAGTCGGAGGGCTGACACCCCTCACTTTCAGAGCCCCACTCGGCCT 240

81 L V A V R R L L V H S S Y H G T T S G  
TGGTGGCTGTGAGGAGGCTCCTGTCCACTCCTCATACCATGGGACCAACCCAGCGGGG 300

101 (D) I A L M E L D S P L Q A S Q F S P I [C]  
ACATTGCCCTGATGGAGCTGGACTCCCCCTTGCAGGCCCTCCAGTTCAGCCCCATCTGCC 360

121 L P G P Q T P L A I G T V [C] W V N G L G  
TCCCAGGACCCAGACCCCCCTCGCCATTTGGACCGTGTGCTGGGTAAACGGGCTGGGG 420

141 V H S G E A L A S V L Q E V A V P L L D  
TCCACTCAGGAGAGGCCCTGGCGAGTGTCTTTCAGGAGGTGGCTGTGCCCCCTCCTGGACT 480

SUBSTITUTE SHEET (RULE 26)

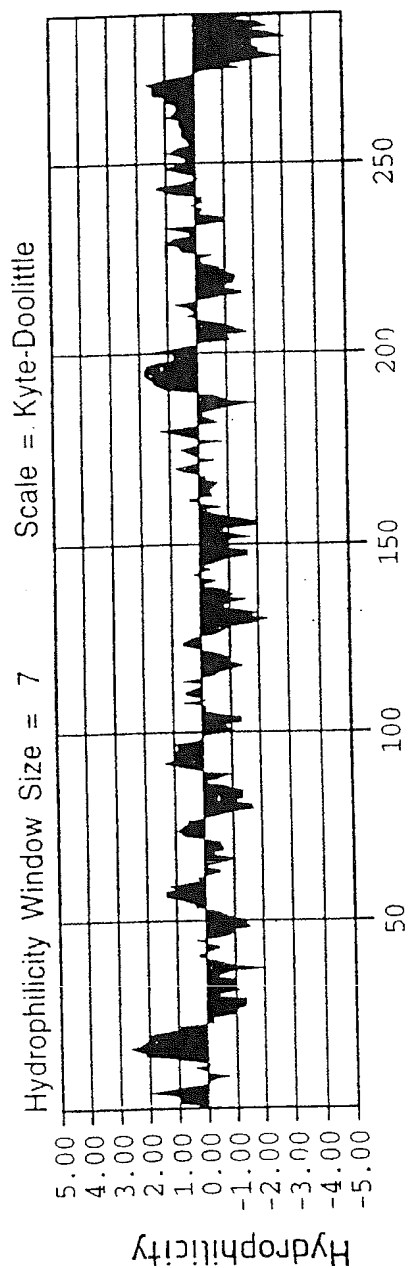
56/62

## FIGURE 20B (All)

CGAACATGTGTGAGCTGATGTACCACTAGGAGAGCCAGCCTGGCTGGCCAGCGCCTCA 540  
161 S N M [C] E L M Y H L G E P S L A G Q R L  
TCCAGGACGACATGCTCTGTGCTGGCTCTGTCCAGGGCAAGAAAGACTCCTGCCAGGGTG 600  
181 I Q D D M L [C] A G S V Q G K K D S [C] Q G  
ACTCCGGGGGGCGCTGGTCTGCCCCCATCAATGATACGTGGATCCAGGCCGGCATTTGTGA 660  
201 D [S] G G P L V [C] P I N D T W I Q A G I V  
GCTGGGGATTTCGGCTGTGCCCCGGCCTTCCGGCCTGGTGTCTACACCCAGGTGCTAAGCT 720  
221 S W G F G [C] A R P F R P G V Y T Q V L S  
ACACAGACTGGATTTCAGAGAACCCCTGGCTGAATCTCACTCAGGCATGTCTGGGGCCGCC 780  
241 Y T D W I Q R T L A E S H S G M S G A R  
CAGGTGCCCCCAGGATCCCACCTCAGGCACCTCCAGATCCCACCCAGTGTGCTGCTTGAGC 840  
261 P G A P G S H S G T S R S H P V L L L E  
TGTGACCGTATGCTTGTGCTTGGTCCCTGTGAACCATGAGCCATGGAGTCCGGGATCCCC 900  
281 L L T V C L L G S L  
TTTCTGGTAGGATTGATGGAATCTAATAATAAA

57/62

FIG 20B (B)



58/62

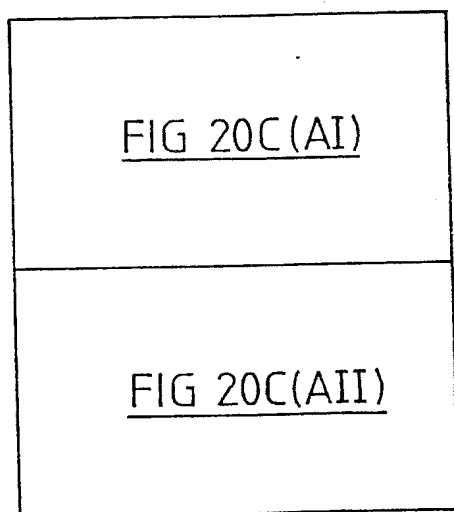


FIG 20C(A)

SUBSTITUTE SHEET (RULE 26)

59/62

**FIGURE 20C (AI)**

CCTGTGTCGCCCCAGGATGCTGAACCGAATGGTGGGGCAGGACACGCAGAGGGCG 60  
 1 [C] G R P R M L N R V M V G G Q D T Q E G  
 AGTGGCCCTGGCAAGTCAGCATCCAGCGCAACGGAAGCCACTTCTGCGGGCAGCCTCA 120  
 21 E W P W Q V S I Q R N G S H F [C] G G S L  
 TCGCGGAGCAGTGGGTCCTGACGGCTGCGCAGTCTCCGCAACACCTCTGAGACGTCCC 180  
 41 I A E Q W V L T A A (H) [C] F R N T S E T S  
 TGTACCAGTCCCTGCTGGGGCAAGGCAGCTAGTGCAGCCGGGACCAACACGCTATGTATG 240  
 61 L Y Q V L L G A R Q L V Q P G P H A M Y  
 CCCGGTGAGGCAGGTGGAGAGCAACCCCTGTACCAGGACGGCCTCCAGCGCTGACG 300  
 81 A R V R Q V E S N P L Y Q G T A S S A (D)  
 TGGCCCTGGTGAGCTGGAGGCACCAGTGCCCTTACCCAATTACATCCTCCCCGTGTGCC 360  
 101 V A L V E L E A P V P F T N Y I L P V [C]  
 TGCCTGACCCCTCGGTGATCTTTGAGACGGGCATGAAGTGGTCACTGGCTGGGGCA 420  
 121 L P D P S V I F E T G M N [C] W V T G W G  
 GCCCAGTGAGGAAGACCTCCTGCCCCGAACCGGGATCCTGCAGAACTCGTGTGCCCA 480  
 141 S P S E E D L L P E P R I L Q K L A V P

SUBSTITUTE SHEET (RULE 26)



60/62

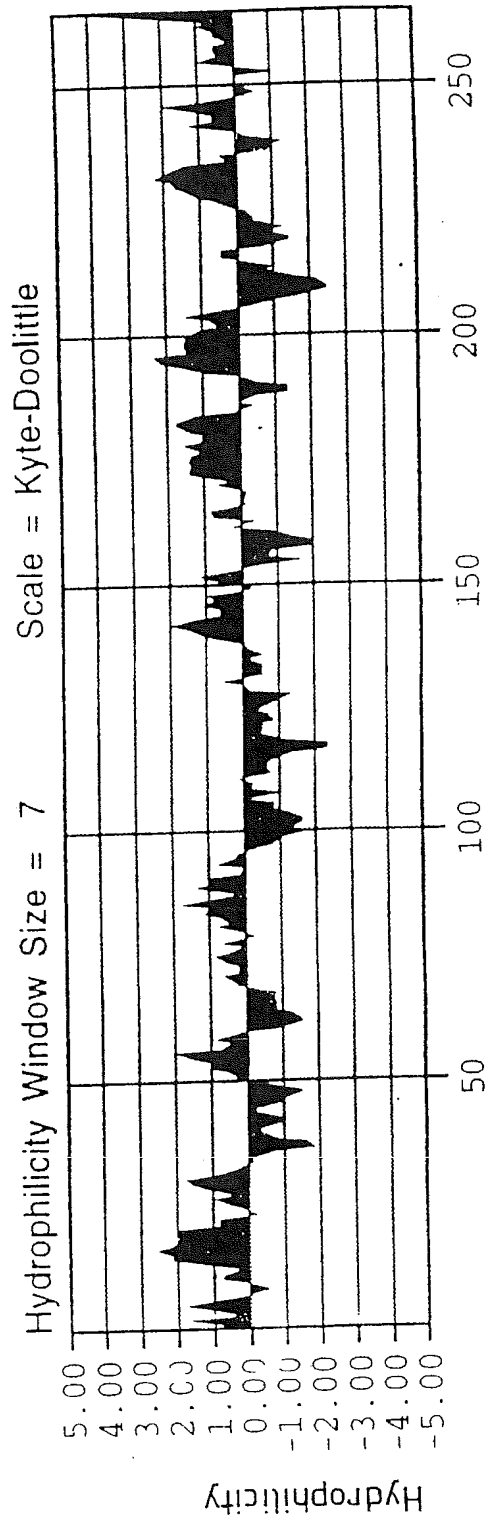
## FIGURE 20C (AII)

TCATCGACACACCCAAAGTGCAACCTGCTCTACAGCAAAAGACACCCGAGTTTGGCTACCAAC 540  
161 I I D T P K [C] N L L Y S K D T E F G Y Q  
CCAAAACCATCAAGAATGACATGCTGTGCGCCGGCTTCGAGGAGGGCAAGAGATGCCT 600  
181 P K T I K N D M L [C] A G F E G K K D A  
GCAAGGCGACTCGGGCGGCCCCCTGGTGTGCTCGTGGTCAAGTCGTGGCTGCAGGCGG 660  
201 [C] K G D (S) G G P L V [C] L V G Q S W L Q A  
GGTGATCAGCTGGGGTGAGGGCTGTGCCCGCCAGAACCCGCCAGGTGTCTACATCCGTG 720  
221 G V I S W G E G [C] A R Q N R P G V Y I R  
TCACCGCCACCAACTGATCCATCGGATCATCCCCAACTGCAGTCCAGCCAGCGA 780  
241 V T A H H N W I H R I I P K L Q F Q P A  
GGTTGGGCGCCAGAAAGTGAGACCCCCCGGGCCAGGAGCCCCCTTGAGCAGAGCTCTGCAC 840  
261 R L G G Q K \* D P R G Q E P L E Q S S A  
CCAGCCTGCCCGCCACACCATCCTGCTGGTCCCTCCAGCGCTGCTGTTGCACCTGTGAG 900  
281 P S L P A H T I L L V L P A L L L H L  
CCCCACGAGACTCATTTGTAAATAGCGCTCCTCCTCCCTCTCAATAACCTTATTTA 960  
TTTATGTTTCTCCCAATAAA

SUBSTITUTE SHEET (RULE 26)

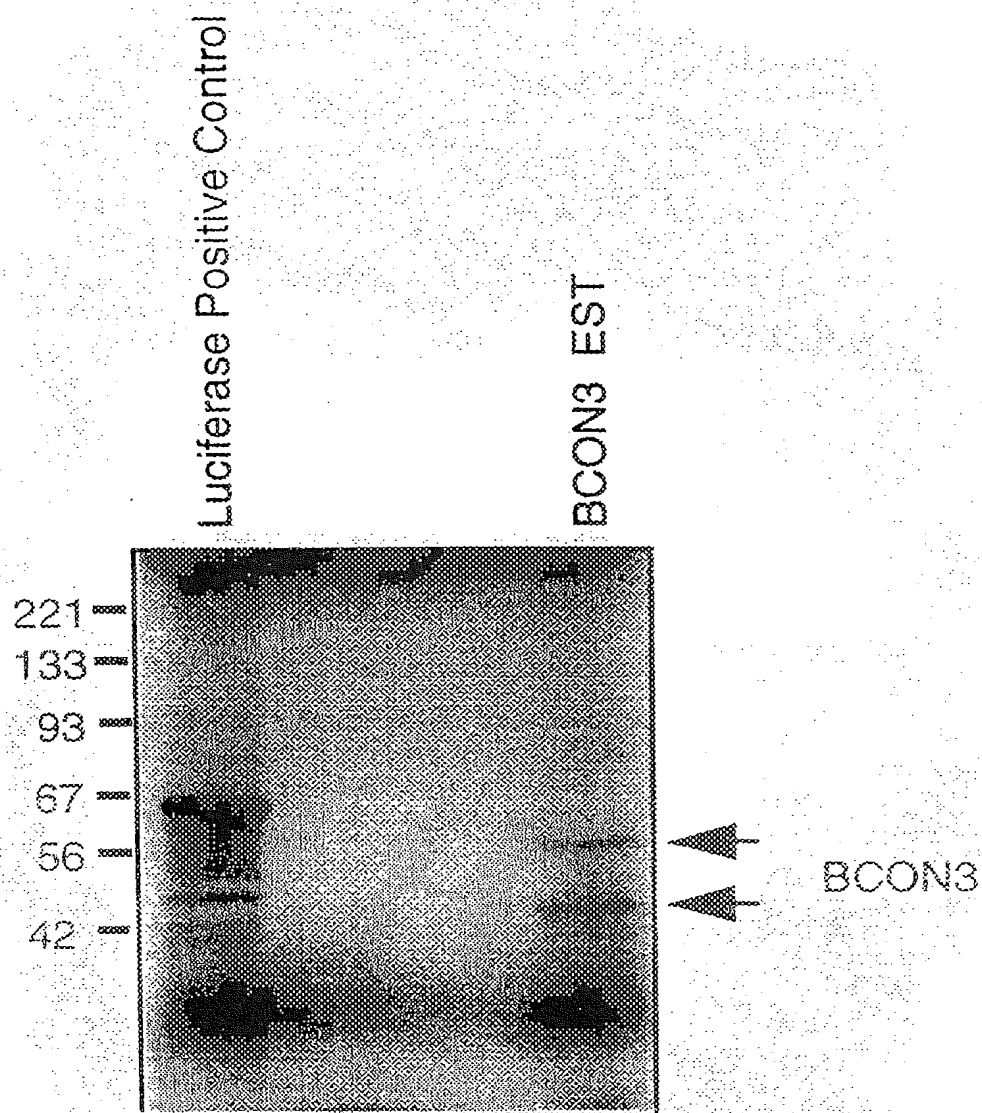
61/62

FIG 20C(B)



SUBSTITUTE SHEET (RULE 26)

62/62

FIG 21

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/AU 98/00085

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
Int Cl <sup>6</sup> : C12N 009/12, 009/64, 015/54, 015/57; C07K 016/40; A61K 038/45, 038/48; C12Q 001/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) STN (DGENE) (see below)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched SEQUENCE DATABASES (see below) MEDLINE (see below)		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) (as online, STN (DGENE): TGGG[AC] [AGT] [GC] T [AGT] AC [AG] GC [AGT] GC [AGT] CA [CT] TG AND GG [AGT] CA [CT] [AT] [GC] [ACT] GG [ACT] CC [ACT] [CT] T and SWISSPROT, GENBANK, EMBL, PIR: SEQ ID Nos: 6, 8 and 10 MEDLINE: 16p13.3 AND "serine protease"		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Proc. Natl. Acad. Sci. USA 87, pages 960-3 (1990) Hanson, R.D. et al. "A cluster of hematopoietic serine protease genes is found on the same chromosomal band as the human $\alpha$ /S T-cell receptor locus." See whole document, especially page 961 column 2-962 column 1.	1,4,7,10,13,16,19,26, 27,31,32,35,38,46,49, 52,56,57,60,63
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input type="checkbox"/> See patent family annex		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 31 March 1998	Date of mailing of the international search report <b>03 APR 1998</b>	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No.: (02) 6285 3929	Authorized officer  <b>JIM CHAN</b> Telephone No.: (02) 6283 2340	

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/AU 98/00085**Box I** Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 19, 20, 26, 31, 44  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
the breadth of the claims was such that it was uneconomical to conduct a search that encompassed the full scope of the claims.
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

**Box II** Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/AU 98/00085

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Proc. Natl. Acad. Sci. USA 87, pages 3811-5 (1990) Vanderslice, P. et al. "Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine protease family." See whole document	1-4,5-7,10-13,14-16,26-28,31,32-38,44,46-52,55-59,60-63
X	J. Reprod. Fertil. 100, pages 567-75 (1994) Bermudez, D. et al. "Proacrosin as a marker of meiotic and post-meiotic germ cell differentiation: quantitative assessment of human spermatogenesis with a monoclonal antibody." See whole document, in particular Introduction.	46-48, 50, 51, 55
X	J. Biol. Chem. 269(29) pages 18843-8 (1994) Yu, J.X. et al. "Prostasin is a novel human serine protease from seminal fluid." See whole document, in particular discussion.	1-3,5,6,10-12,14,15,27,28,31-34,37,44,46-48,50,51,55-59,61,62
X	J. Biol. Chem. 269 (31) pages 19976-82 (1994) Matsushima, M. et al. "Structural characterisation of porcine enteropeptidase." See whole document, especially figure 4.	1-3,5,6,10-12,14,15,27,28,31-34,37,44,46-48,50,51,55-59,61,62
X	J. Biol. Chem. 270 (22) pages 13483-89 (1995) Yu, J.X. et al. "Molecular cloning, tissue-specific expression, and cellular localisation of human prostasin mRNA." See whole document, especially figure 2, Introduction and Discussion.	1-3,5,6,10-12,14,15,27,28,31-34,36,37,41-43,44,46-48,50,51,55-59,61,62
X	Mol. Reprod. Dev. 43, pages 236-47 (1996) O'Brien, D.A. et al. "Boar proacrosin" expressed in spermatids of transgenic mice does not reach the acrosome and disrupts spermatogenesis." See whole document.	1-3,5,6,10-12,14,15,27,28,31-34,36,37,41-43,44,46-48,50,51,55-59,61,62